Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

THE OHIO STATE UNIVERSITY CRERAR LIBRARY DEC 18 1952

CONTENTS

Need and Perceptual Change in Need-Related Objects: J. C. GILCHRIST AND L. S. NESBERG 369
The Stimulus Trace in Effectors and Its Relation to Judgment Responses: R. C. DAVIS ... 377
The Projective Measurement of Experimentally Induced Levels of Sexual Motivation: R. A. CLARK 391
Effect of Incentives upon Reactions to Peripheral Stimuli: H. P. BAHRICK, P. M. FITTS, AND R. E. RANKIN 400
Placing Precision and Angle of Regard: A. J. RIOPELLE AND T. M. STRITCH 407
Effect of Discrimination Reversal on Human Discrimination Learning: R. D. WALK 410
Reminiscence in a Manipulative Task as a Function of Work-Surface Height, Prerest Practice, and Interpolated Rest: D. S. ELLIS, V. MONTGOMERY, AND B. J. UNDERWOOD 420
Failures of Attention in Selective Listening: D. E. BROADBENT 428
Situational Instructions and Task Order in Recall for Complete and Interrupted Tasks: D. G. HAYS 434
The Effect of Long Intertrial Intervals on the Limit of Eyelid Conditioning: M. R. BARON 438
Individual Differences in Defensive Forgetting: C. W. ERIKSEN 442
Temporal Numerosity: I. Perceived Number as a Function of Flash Number and Rate: P. G. CHEATHAM AND C. T. WHITE 447
Figure-Ground Contrast and Binocular Rivalry: L. T. ALEXANDER AND P. D. BRICKER 452
Discrimination of Tactual Stimuli: H. J. BAUER 455
Visual Factors in the Perception of Verticality: C. W. MANN 460
Erratum: C. E. NOBLE 465

American Psychological Association

Vol. 44 No. 6 December 1952

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.00, or $14.00 annually. Single copies are $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free reprints without covers. All of the cost of an author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas. Priority in publication is given to articles whose authors assume the full cost of publication.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 44, NO. 6 DECEMBER, 1952

NEED AND PERCEPTUAL CHANGE IN NEED-RELATED OBJECTS [1]

J. C. GILCHRIST AND LLOYD S. NESBERG

University of Wisconsin

Such recent studies as have dealt with the distorting effects of need upon the perception of objects relevant to that need (1, 2, 3, 5, 6) have been difficult to interpret unequivocally as some of the investigators cited have pointed out. The major difficulties have been methodological and are two in number.

1. The investigations cited have used standard psychophysical methods in which the standard and variable stimuli have differed in dimensions other than that in which *S*s were to make their stimulus matches. This fact leaves open the possibility that any effects found may be due solely to the psychophysical relationships between the two stimuli. In attempting to control the above obvious flaw, the use of a range of need-related objects having different reward value and/or the use of control, non-need-related objects has been introduced. Comparisons have then been made within the range of need-related objects and/or between that range and the control set. This technique merely removes the difficulty one step from being readily apparent. In order to use either or both of these control techniques it has still been necessary to make the standard stimuli differ in some dimensions other than that of need relatedness. If the effects of these differences (e.g., those involved in patterning of impressed design) were known, they could be removed from the result. But it seems fair to say that not even the number, much less all the effects, of these factors are known and, therefore, any differences obtained cannot be unequivocally attributed to the need relatedness of the objects.

2. The above statements only apply to the reward-value aspect of need relatedness, for if it can be shown under two or more conditions of need that perception of objects changes as a function of that need, then such changes may be attributed to need. But here, too, the studies under consideration have been somewhat less than clear-cut.

The "economic" need with which most investigators have worked has been only tenuously inferable. For example, to infer that children of low-income families need money more than do the children of families having a relatively high income, means to assume that the children of both groups have taken over the norms and aspirations further assumed to be those of adult, middle-class American society, and that the "poor" children have a greater negative attainment discrepancy between their economic performance and aspirations than do the rich children, or that they experience what discrepancy exists more keenly. If one or more of these assumptions is unverified, the attributing of positive results to need alone is difficult.[2]

[1] This study was supported in part by funds provided by the Research Committee, The Graduate School, University of Wisconsin.

[2] That the same general line of argument can be applied to studies using swastikas and dollar

It was in view of the above criticisms that the present experiments were undertaken. Clearly in terms of the first major criticism of the previous experiments, control objects or a differing standard and variable cannot be used. This situation means, of course, that a time difference between the presentation of the standard and the variable must be used since, if they are to differ only in the dimension in which the stimulus match is to be made, simultaneous presentation of identical objects would result in both objects being subject to the same distortion effects, if any, and, therefore, such matches as would be made would not differ in their "errors" from those found in the classical psychophysical experiment. As to the second criticism, the obvious answer is to make use of some motivational component generally recognized as inferable in terms of a direct measurement of time of deprivation, e.g., hunger or thirst. The present experiments, then, are ones which might best be described as attempts to show the effects of need upon the short term perceptual recall of need-related objects: the effect, more specifically, of progressive deprivation upon the illuminance matches of projected color pictures of need-related objects.[3]

Experiment I

Procedure.—The apparatus for the present experiment consisted of a 2 × 2-in. slide projector (The Three Dimension Co. Vivid Model "D") with a 300-w. projection lamp modified so that the cooling fan unit operated on a separate circuit from that of the lamp. Into the lamp circuit was inserted a General Radio Co. Variac Model VIOH arranged to permit the lamp filament voltage to be varied from 0–130 v. Actual lamp voltage was read from a voltmeter whose characteristics permitted readings accurate within ±.5 v. The projector was placed on a stand to the rear and above *S*'s chair and focused upon a beaded projection screen 12 ft. to the front of *S*'s chair. To the right of *S*, the Variac, screened so that it could be reached but not seen by *S*, was placed on a low laboratory table. The *E* stood to the right rear of *S* where he could manipulate the slide projector and Variac without being observed by *S*. The stimulus objects were four color slides of food objects. They were complex pictures of complete or almost complete meals of the type usually found in illustrated homemakers' magazines from which they were reproduced. The main food items of the four pictures were (*a*) T-bone steaks, (*b*) fried chicken, (*c*) hamburgers, and (*d*) spaghetti.

During a given session, the procedure was as follows: *S* was seated in the already darkened room and told: "This is an experiment in matching pictures. I will project a picture on the screen there (points) and you are to look at it for 15 sec. At the end of that time, I will turn the picture off and then a few seconds later I will turn it back on. When it is turned back on, however, it will not look the same. *Your job is to make it look the same as before* by turning this knob here (has *S* put hand under the curtain and feel the Variac knob). You may turn the knob only in one direction on each match. If you go past the setting you want, let me know and we will start over." After the instructions, the slides were presented in random order twice at each of two standard lamp voltages, 60 and 90 v.,[4] once in which the match was made from a voltage lower than the standard and once from a higher voltage. The presentation was made by (*a*) closing the projector aperture by a slide-

signs with native and refugee *S*s has been pointed out by Klein, Schlesinger, and Meister (5).

[3] Illuminance matches were chosen as the dependent variable rather than the usual size matches, since previous work (5, 6) has led the present authors tentatively to accept the hypothesis that young normal adults (who were the *S*s in the present experiment) have become able to dissociate size-judging behavior from the value aspects of the object being judged.

[4] Since the slides were complex in design as well as varying both between and within slides as to hue, saturation, and brightness, presenting the data in illuminance units would seem to add nothing but an unwarranted appearance of elegance to the results. Our apologies are extended to all purists in the field of visual research for our use of the term "illuminance" when the measurements were actually made in terms of lamp voltage which does not vary rectilinearly with it. Without slides the projection system yielded illuminance values at the center of the surface of the screen of 3.7 ft.-candles at 60 v. and 15.0 ft.-candles at 90 v., and, between 60 v. and 120 v., illuminance values and lamp voltage do have an essentially rectilinear relationship.

carrier shutter, (*b*) setting lamp voltage at the standard, (*c*) opening the shutter for 15 sec., (*d*) closing the shutter for 10 sec. during which time the lamp voltage was reset at either plus or minus 20 v. of the standard, and (*e*) opening the shutter until *S* had completed his match and repeating for next presentation until all slides had been presented at both standards at both the "in" and "out" lamp voltage settings.

The *Ss* were 26 University of Wisconsin students in an elementary psychology course, all of whom volunteered to go 20 hr. without food and all liquids except water. Each *S* was assigned at random to one of two groups. The first of these groups, to be referred to as the hungry group, went through one matching session right after their noon meal, one session right after the time they usually would have eaten the evening meal that same day and a final session right after the time they usually would have eaten breakfast the next morning. They were thus 0, 6, and 20 hr. food deprived at the three sessions, respectively. The second group, the satiated group, followed the same routine except that they ate each of the two meals missed by the hungry group.

Results.—The experimental design permitted an analysis of variance with the following main effects: slides, sessions, hunger, illuminance of standard (included only so that possible interactions with the other variables could be evaluated) and *Ss* within hunger groups. Since neither slides nor any of the interactions of slides or illuminance proved significant, these were combined to yield the breakdown given in Table 1. The p values for the reported analysis are of the same magnitude as were found in the longer analysis.

TABLE 1

ANALYSIS OF VARIANCE: EXPERIMENT I

Source	*df*	Mean Square	*F*	p
Hunger groups	1	5703.7	1.82*	
Sessions	2	2092.4	4.23	<.05
Ss within hunger groups	24	3140.5	6.35	<.001
Hunger groups × sessions	2	3081.7	6.23	<.01
Residual	48	494.7		

* Tested by *Ss* within hunger groups.

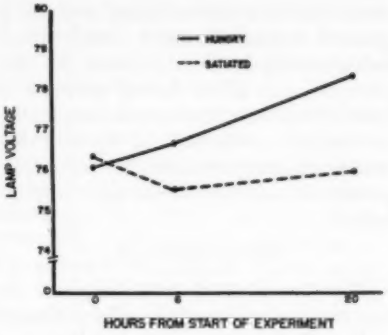


FIG. 1. Mean lamp voltage matches of the hungry and satiated groups as a function of time from the start of Exp. I

The point at issue in the experiment is, of course, the interaction between hunger groups and sessions. Table 1 shows the interaction to be significant and Fig. 1 is a plot of the mean lamp voltage of the two groups at each of the three sessions. An accurate mean match would have occurred with mean lamp voltage settings of 75 v.

Discussion.—It is readily seen from Fig. 1 that the progressively hungrier *Ss* show a directly increasing positive time error while the satiated *Ss* show a positive time error which remains fairly stable throughout the three sessions. Although the data are as yet too meager to make any definite commitments, they seem to warrant suggesting that the motivational effects of need of *S* and goal relatedness of the stimuli being judged as well as background and anchoring stimuli must be taken into account in the formulation of such concepts as Helson's adaptation level (**4**). There remain, however, issues which must be settled before that step can be taken or even before the issues of need relatedness of the stimulus can be settled: (*a*) The experiment dealt with hunger and illuminance matches of food objects, but there is no guarantee in the

data that the effects found are not a general visual function rather than being confined to food objects. (*b*) The motivational effect found may be a purely local one, restricted completely to hunger. Because of these difficulties of interpretation, the other experiments reported below were undertaken.

Experiment II

This experiment was designed to test the generality of the effects found in Exp. I: It is an attempt to demonstrate the same phenomena using thirst and thirst-related objects.

Procedure.—The apparatus for this experiment is the same as that used in Exp. I except for the projection slides. The slides in this case were three glass-bound Kodachrome slides, each photographed with the same dark-blue background. In each photograph was a pitcher and glass as the only objects. One of the pitcher and glass sets contained ice water, another orange juice, and the third milk.

During a given matching session, the procedure duplicated that of Exp. I except that standard lamp voltages here were 60 v. and 100 v.

The *Ss* were 48 students in the beginning courses in psychology or in sociology at the University of Wisconsin. All *Ss* had had less than one full semester of psychology and all volunteered to go at least 8 hr. without food or liquids of any kind. Each *S* had four experimental sessions separated by 2½ hr. ± ½ hr. Each *S* had, therefore, sessions at 0, 2½, 5, and 7½ hr. after the start of the experiment and each session required approximately ½ hr. There were four groups of randomly assigned *Ss*. Groups I and II were allowed to eat and drink normally and before each session obtained a drink of water at a bubbler outside the laboratory. The difference which existed in their treatment was that for Group I: the same pitchers and glasses filled with the same liquids as appeared in the slides were on the table to their left, while for Group II the pitchers, etc. were absent. Groups III and IV were the counterparts of I and II respectively except for the fact that they did not eat or drink after the beginning of the first session. There was, therefore, Group I, which was satiated with goal objects present; Group II, satiated with goal objects absent; Group III, progressively thirstier with goal objects present; and Group IV, progressively thirstier with goal objects absent. Groups III and IV were given, at the end of the fourth session, as much ice water, orange juice, and/or milk, as they wished to drink and then, without any forewarning, were immediately asked if they would again be seated for a fifth matching session.

Results.—Separate analyses of variance were run to test the effects of the presence of the pitchers, etc., on the table to *S*'s left. Since this goal-object presence proved to have no effects,[5] Groups I and II and Groups III and IV were combined to provide two groups whose only difference as groups lay in the degree of deprivation through the four matching sessions they had in common. Illumination levels of the standard and the separate slides were also combined to produce an analysis of variance, the results of which are given in Table 2.

TABLE 2

Analysis of Variance: Experiment II

Source	*df*	Mean Square	*F*	*p*
Thirst groups	1	.38	—*	
Sessions	3	34.36	42.95	<.001
Ss within thirst groups	46	8.11	10.14	<.001
Thirst groups × sessions	3	55.89	69.86	<.001
Residual	138	.80		

* Tested by *Ss* within thirst groups.

The point to be considered is the significant interactions between thirst groups and sessions. Figure 2 is a plot of the mean illumination matches of the sessions of the groups. $p < .001$ for the difference between the means of the fourth and fifth (after drinking) sessions of the thirsty *Ss*[6] while

[5] The only reason for including this variable in the experiment at all was a methodological one. It was necessary to know its effect, if any, before designing an experiment planned for a later date.

[6] All *Ss* drank at least 8 oz. of one of the three liquids. The median amount consumed was 16 oz. and the most frequently chosen liquid was orange juice. Only three *Ss* chose milk.

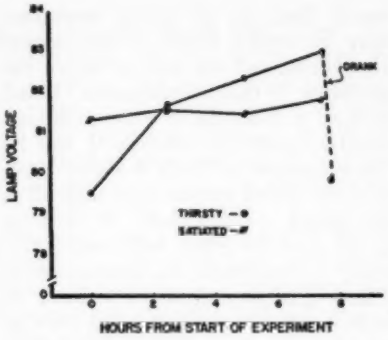


FIG. 2. Mean lamp voltage matches of the thirsty and satiated groups as a function of time from the start of Exp. II

.10>p>.05 for the difference in the means of their first and fifth sessions.

Discussion.—As in the previous experiment, increasing need gave rise to an increasingly positive time error in the illuminance matches of need-related objects.[7] The fact that this effect occurred with a different need constellation and different need-related objects would appear to support the more general statement of the effect than the previous findings would permit. It might be argued, however, that since *S*s were also hungry as well as thirsty in the present experiment, it is nothing more than a repetition of the previous one with the only change being the kind of need objects present. The argument has its cogent aspects, but it is believed that the effects found herein are primarily attributable to thirst: first, the total amount of change found in 8 hr. in the present experiment with hunger and thirst is greater than that previously found in 20 hr. with hunger alone; second, most of the effect disappeared immediately after *S*s drank and since 21 of them drank only ice water or orange juice, they were probably still hungry.

Although the correctness of the general statement "Increasing need gives rise to increasingly positive time errors in the illuminance matches of objects relevant to that need" now seems more probable, a major difficulty remains. The necessity of the goal objects being relevant is not established. Clearly, the effect does occur when matches are made of goal objects relevant to the need, but does it also occur when the objects to be matched are not relevant?

EXPERIMENT III

One aspect of the question just posed relates to whether or not any objects (figures or shapes) are necessary for the effect to occur. That is to say, is the effect a general visual function occurring regardless of the structure of the stimulus field? The simplest case by which this question may be approached is that of homogeneous stimulus areas. The present experiment, then, compares the effect of deprivation upon colored, homogeneous stimulus areas and stimulus areas containing need-related objects.

Procedure.—The apparatus for this experiment is the same as that used in Exp. II except for the addition of six slides. These slides presented *S*s with homogeneous color fields of red, orange, yellow, green, blue-green, and blue having the same area as that given by the need-related slides.

During a given matching session, the procedure here duplicated that of the previous experiments except that only one standard lamp voltage was used: 80 v.

The *S*s were 48 students having normal color vision in the beginning course in psychology at the University of Wisconsin all of whom volunteered to go 8 hr. without food or liquids of any kind. Each *S* underwent four experimental sessions the first three of which were separated by 4 hr. ± ¼ hr. The fourth session followed the third by about 5 min. in which interval each *S* drank at least 8 oz. of ice water. There were

[7] That the progressively thirstier *S*s started far below the satiated *S*s and then increased their matches but never as far above as they were below, should not disturb this interpretation when it is remembered that the between-group effects are independent and the between-sessions effects correlated.

TABLE 3

ANALYSIS OF VARIANCE FOR GROUPS I AND III: EXPERIMENT III

Source	*df*	Mean Square	*F*	*p*
Thirst groups	1	18.71	1.04*	
Sessions	2	9.95	2.99	
Ss within thirst groups	22	18.04	5.42	<.001
Thirst groups × sessions	2	2.46	—	
Residual	44	3.33		

* Tested by *Ss* within thirst groups.

four groups of randomly assigned *Ss*. Groups I and II were allowed to eat and drink normally and before each session obtained a drink of water at a bubbler outside the laboratory. The difference in their treatment was that Group I matched only the six homogeneous color slides and Group II matched the three need-related slides and three of the homogeneous color slides, the orange, green, and blue. Groups III and IV were the counterparts of I and II respectively except for the fact that neither ate nor drank from the first to third sessions.[8]

Results.—The main comparisons to be made are for the first three sessions of Groups I and III and Groups II and IV. The results for Groups I and III are given in Table 3 where it will be noted that, for all slides combined, only *Ss* within thirst groups proved significant, and the term of major importance to this investigation, Thirst groups × Sessions, is not significant. Similar results for Groups II and IV are presented in Table 4 where kinds of slides (color versus need-related) is an added main effect. It will be noted that the effect with which this study is concerned, the interaction of Thirst groups × Sessions × Kind of slide is significant at the .1% level of confidence. In order to locate the source of the significance of this term, analyses of variance were performed for each group separately. In the analysis of Group II, only individual differences terms proved significant. For Group IV, however, those terms plus the Sessions × Kind of slide interaction were significant with the latter being associated with a $p < .001$. From these results, it is clear that the significant effect was due to the progressively thirstier group (IV) acting differently on the two kinds of slides. Figure 3 is a plot of that group's mean matches of the two kinds of slides by sessions.

The remaining results to be presented are those of the differences between the third and fourth sessions. These are given in Table 5.

[8] The reason for the different slide treatments lay in the fact that we did not know what the effects of mixing kinds of slides within a group might be.

TABLE 4

ANALYSIS OF VARIANCE FOR GROUPS II AND IV: EXPERIMENT III

Source	*df*	Mean Square	*F*	*p*
Thirst groups	1	.77	—*	
Sessions	2	11.86	3.10**	
Kind of slide	1	90.73	4.30***	<.05
Ss within thirst groups	22	37.58	22.11***	<.001
Thirst groups × sessions	2	1.36	—	
Thirst groups × kind of slide	1	10.83	—***	
Sessions × kind of slide	2	4.28	2.52	
Sessions × *Ss* within thirst groups	44	3.83	2.25	<.01
Kind of slide × *Ss* within thirst groups	22	21.11	12.42	<.001
Thirst groups × sessions × kind of slide	2	27.46	16.16	<.001
Residual	44	1.70		

* Tested by *Ss* within thirst groups.
** Tested by sessions × *Ss* within thirst groups.
*** Tested by kind of slide × *Ss* within thirst groups.

Discussion.—It is clear from the above data that the effects under investigation are not a general visual function: that some degree of need relatedness must obtain in the to-be-matched stimuli before the increasingly positive time error as a function of deprivation is found. It is clear, that is, unless an argument be advanced to the effect that such change may be found in the presence of any articulated stimulus but not with a homogeneous field.

TABLE 5

MEAN DIFFERENCES BETWEEN THE BEFORE AND AFTER DRINKING SESSIONS: EXPERIMENT III

Group	Mean Diff. (Volts)*	t	p
I	+0.21	0.28	>.10
III	−1.18	2.04	>.05
II			
Need-related slides	−2.06	5.02	<.001
Color slides	+0.91	1.66	>.10
IV			
Need-related slides	−0.09	0.22	>.10
Color slides	−1.15	1.98	>.05

* A plus sign means that the fourth session match was higher than that of the third; a minus sign that it was lower.

EXPERIMENT IV

Any final answer to the argument that the results reported so far in this investigation are due to articulation of the stimulus field rather than need is, in our present state of knowledge, impossible. No matter which set of articulated non-need-related objects were chosen as control objects, it could always be argued that any differential effects obtained between them and the need-related objects were due to the particular form qualities of the two sets of objects. Nevertheless, by the use of control stimuli involving forms, it is possible to require that possible proponents of such a position specify the conditions under which the hypothesis may be tested. Such is the purpose of the present experiment.

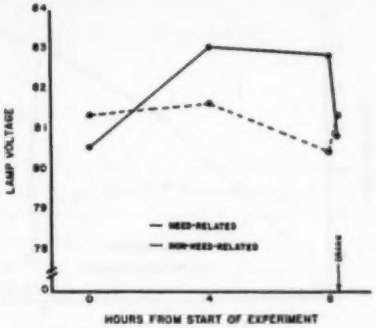


FIG. 3. Mean lamp voltage matches of need- and non-need-related slides by a progressively thirstier group (Group II) as a function of time from the start of Exp. III

Procedure.—In view of the results of Exp. III, it was decided that only one group of *Ss* who neither ate nor drank for 8 hr. and who matched both sets of slides was required. Thus the treatment of *Ss* in the present experiment is analogous to that of Group III of the Exp. III except that the control slides were color landscape slides of (*a*) a range of mountains, (*b*) a grove of trees in autumn, and (*c*) a park lawn and shrubs.

The *Ss* were 12 students in the beginning course in psychology at the University of Wisconsin all of whom volunteered to go 8 hr. without food or liquids of any kind.

Results.—The main results of the experiment are presented in Table 6 and Fig. 4. It will be noted there

TABLE 6

ANALYSIS OF VARIANCE FOR THE FIRST THREE SESSIONS OF EXPERIMENT IV

Source	df	Mean Square	F	p
Sessions	2	33.30	—*	
Kinds of slides	1	40.05	1.16*	
Ss	11	31.42	9.64	<.001
Sessions × kinds of slides	2	34.66	10.62	<.001
Sessions × *Ss*	22	3.85	1.18	
Kinds of slides × *Ss*	11	2.05	—	
Residual	22	3.26		

* Tested by sessions × kinds of slides.

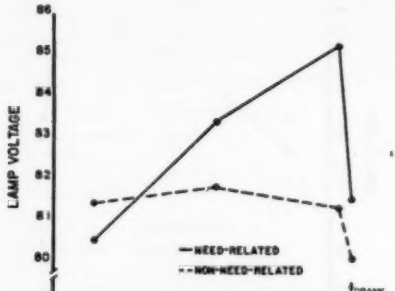


FIG. 4. Mean lamp voltage matches of need- and non-need-related slides by progressively thirstier *Ss* as a function of time from the start of Exp. IV

that the Sessions × Kind of slide interaction is significant ($p < .001$) and that the curve forms are essentially the same as that for Group III of Exp. III. The *t* values associated with the changes between the third and fourth sessions are (*a*) for need-related slides significant at the .1% level of confidence, and (*b*) for non-need-related not significant ($p > .05$).

Discussion.—On the basis of these results, it would appear that it is not just any articulation of the stimulus field which produces increasingly positive time errors with increasing need. And that, if an articulation position is to be advanced to explain these results, it must be on more specific, testable grounds than "articulation" alone. Taken together with the previous three experiments, the present findings must be interpreted as favoring the position that the effect of need upon the time error is restricted to need-related objects.

Statements as to time error and need relatedness, however, require two more lines of investigation before they may be made with impunity. These are (*a*) a systematic study of the development of such time error effects, and (*b*) an inquiry into the effects of wider degrees of need relatedness.

Summary and Conclusions

A series of experiments were performed in the area of need and perceptual change in need-related objects. Specifically, the effects of hunger and thirst upon the illuminance matches of projected pictures of food-related and liquid-related objects were investigated.

The statement "Increasing need gives rise to an increasingly positive time error in the illuminance matches of objects relevant to that need" is supported by the evidence presented.

(Received June 16, 1952)

References

1. Bruner, J. S., & Goodman, C. C. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, 1947, **42**, 33–44.
2. Bruner, J. S., & Postman, L. Symbolic value as an organizing factor in perception. *J. soc. Psychol.*, 1948, **27**, 203–208.
3. Carter, L. F., & Schooler, K. Value, need and other factors in perception. *Psychol. Rev.*, 1949, **56**, 200–207.
4. Helson, H. Adaptation-level as frame of reference for prediction of psychophysical data. *Amer. J. Psychol.*, 1947, **60**, 1–29.
5. Klein, G. S., Schlesinger, H. J., & Meister, D. E. The effect of personal values on perception: an experimental critique. *Psychol. Rev.*, 1951, **58**, 96–112.
6. Lysak, W., & Gilchrist, J. C. Value, equivocality and goal availability as determinants of size judgments. *J. Pers.*, in press.

THE STIMULUS TRACE IN EFFECTORS AND ITS RELATION TO JUDGMENT RESPONSES

R. C. DAVIS

Indiana University

This experiment was planned to find, by recording action potentials, whether there is an excitation in the skeletal muscles corresponding to the "stimulus trace"; and, if there is, what its determining factors are; then to discover what relation it has to judgment responses given during its existence.

"Stimulus trace" is ordinarily an inference made to explain data from conditioned response and successive comparison experiments; but the meaning in this report is simply "some process aroused by a stimulus and lasting a few seconds after it." Such a process will be of special interest when it is not the accompaniment of any overt movement.

Stimulus-trace theories generally postulate a process which is initiated by the stimulus and follows a regular fixed course of arousal and decline. With some reason to believe that muscular action will show an excitation concurrent with an auditory stimulus (3, 4, 5), a question in the present experiment was whether such stimulus effects might vary with the situation instead of maintaining the same temporal pattern. The stimulus trace, as a hypothetical construct, is supposed also to have certain relations to an ensuing response. The independent determination of a trace should make possible the exploration of this relation.

The general situation used in this experiment was the familiar one in which *S* is required to judge the intensity of auditory stimuli given successively in pairs. Certain particular features, the intrapair interval, the duration of the first member of the pair and the instructions to *S*, were varied so that the muscular excitation produced by the first stimulus in each pair could be observed under different conditions, and so that, as classic results suggest, the proportions of judgments would vary.

Three previous experiments are especially pertinent to the first problem: the independent measurement of the stimulus trace. Peak (14) compared the "time error" (the basis of trace theory) in judgments with that in eyelid reflex responses to the same stimuli. She found a good deal of correspondence. A period of inhibition of the eyewink caused by prior response corresponds to a period of negative time error, and the two vary together with a number of other determining factors. Peak concluded that both depend on the same basic excitation process to a large extent. In her studies, there is the closest approach to a physiological measurement of the theoretical trace.

Two studies have used muscular action potentials at the site of stimulation in the study of the trace effects in the lifting of weights. Freeman and Sharp (6), with electrodes attached to the muscle doing most of the work, recorded the development and decay of activity produced by the lift. They found the decay was cyclic, corresponding generally to the expected course of the time error when judgments of weight pairs are taken. Payne and Davis (13) used a fixed interval between the members of a weight pair and explored the relation between the judgments and the level of potential in the active arm.

A number of studies using no physiological measure show a relation between the judgments of the second stimulus and the interval between the two stimuli (variation in the "time error"). Postman's (15) study, with very careful techniques, confirms the time-error relationship. Harris (7), whose data show the same general trend, points out the variability in them, a variability which means there must be undiscovered factors in the relation. Later Postman and Page (16) found much resemblance between time-error and retroactive-inhibition effects. Early in the study of time-error phenomenon Köhler (10) theorized that the first stimulus aroused a cortical excitation which leaves behind it a "trace" first rising, then falling as time passes. The dif-

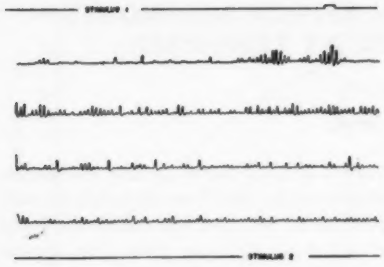


FIG. 1. Section of record showing action potentials integrated over .1-sec. intervals. Response lines are, from top to bottom: right arm, left arm, right masseter, left masseter. Stimuli are indicated by interruption in the signal lines and *S*'s key press by the deflection of the signal line seen in upper right.

ference between the second excitation and the remains of the first, he conceived, is the immediate basis of judgment. In any case, in planning the present experiment it seemed that one could count on getting a systematic variation in the per cent of stronger and weaker judgments by injecting the second stimulus at various intervals after the first, and possibly by varying other conditions. Muscular excitations, recorded by means of action potentials, could then be compared with these.

In the present experiment, the plan was to secure a record of excitation from the site of the response during the stimulus sequence. Accordingly, *S*s were required to respond by key pressing with the right or left hand to signal stronger and weaker comparison stimuli. Electrodes were placed so as to record the activity level of muscles largely concerned in the key-pressing responses, a separate recording system being used for each arm. Two more recording systems were used to record the activity of the right and left masseters to give a suggestion of what might be going on in other segments of the musculature.

APPARATUS

The apparatus used for recording consisted of four channels of high-gain amplification. Signals from these were rectified and integrated electronically (2). At the end of each .1 second the integrators were reset to zero by a special electronic circuit. The result was a series of sawtooth waves, the height of each representing the sum of the action potentials during .1 sec. These were recorded, after suitable power amplification, with a mechanical oscillograph (Electro-Medical Laboratories) using "Teledeltos" paper (see Fig. 1). In recording, the over-all gain was adjusted according to requirements for each *S*, and the amount of gain was ascertained by making records of a known input voltage during each sitting so that the reading could be converted to microvolts.

Stimuli used were 1000-cycle tones; their duration and intrapair interval was controlled by an electronic device containing a number of "flip-flop" circuits, following the design described by Roush (17) but arranged to deliver only one sequence at a time upon the closure of a manual switch. The presence of the stimuli was signalled on the record by having this timer also turn off the writing voltage on one of the oscillograph styli when a tone was being delivered. When *S* gave his judgment by pressing one or another of a pair of keys, this was indicated on the record by a deflection of one or the other of these signal marking styli (Fig. 1).

PROCEDURE

Active electrodes consisting of 5-mm. metal dishes were attached to the dorsal surface of each forearm 2 in. below the elbow where the movement of the extensor digitalis can be palpated (certain exceptions to this will be noted). The reference electrodes for these were saline-soaked felt pads located on the upper arm. An active electrode was attached over the crest of each masseter, and a common reference electrode for these was a clip attached to the left earlobe. The *S* was told that pairs of tones would be delivered from time to time, and that he was to lie on his back as relaxed as possible on the cot with the forefinger of each hand resting lightly on one of the two keys that lay beside him. He was then given instructions specific to the particular experiment. Leads were attached, and *S* was left alone in his separate room with the understanding that he would be given about 5 min. to rest before the stimuli began. During this time *E* watched the oscilloscopes, and if they indicated an unusually high level of activity in some location, he made specific suggestions for further relaxation.

All tones delivered may be considered as equal in intensity. Half of the time, in irregular sequence, the intensity of the second tone of each pair was a fraction of a decibel less than the first, but this difference was so small it produced

no detectable effect on the judgments made by *Ss* nor on their action potentials. Pairs of stimuli were delivered at approximately 1-min. intervals, with each *S* getting 16 pairs.

The *Ss* were students from elementary classes in psychology. The particulars of each procedure are as follows:

Condition 1: Key pressing without intensity discrimination.—The 16 *Ss* were instructed simply to press a designated key at the conclusion of each stimulus pair. Half of them were told to press always with the preferred hand. All stimuli were of 2-sec. duration and the interval between the members of a pair was always 5 sec.

Condition 2: Intensity judgments following various intrapair intervals.—The *Ss* were instructed to press one key if the second tone was stronger than the first, the other if it was weaker, and to guess if uncertain. For half the *Ss* the preferred hand meant stronger, for half it meant weaker. This condition was run first with 16 *Ss*, and then 16 more were run at a later date with the difference that the forearm electrodes were placed slightly toward the inside of the arm rather than on the dorsal surface, producing a level of action potential definitely lower. The data, however, will be treated together. (In the analysis of variance, this change of location was treated as an additional variable and its effects removed from the error term with the loss of a degree of freedom. Comparisons between the absolute levels for the different conditions were not an issue in the experiment.) Stimuli were 2 sec. in duration. The intervals between the members of a pair were .1, 1.0, 3, or 5 sec. and were used in irregular but counterbalanced order in four different sequences for different *Ss*. For the group, each appeared an equal number of times in each serial position.

Condition 3: Intensity judgment following a constant intrapair interval.—Instructions were the same as for Cond. 2, but all intervals between the members of pairs were 5 sec. The electrodes were always on the inside of the forearm. Eight *Ss* were used.

Condition 4: Intensity judgment with constant intrapair interval and short first stimulus.—The procedure was like that for Cond. 3 except that the first stimulus of each pair lasted for 1 sec. Eight *Ss* were used, with records from the arms only.

Measurement and transformation of data.—It was necessary to use a sampling procedure in measuring records. Each run was divided into seconds, the onset of the first stimulus being taken as the origin. Then in each second the largest peak was measured by use of scales marked in microvolts. About 35,000 measurements were made.

Previous studies (**1, 3**) had indicated that: (*a*) muscular responses to auditory stimuli are roughly proportional to the base level of muscular activity, at least under certain conditions, and (*b*) the distribution of the responses is quite skewed. A confirmation of this skewness was found in the present data by tabulating a number of distributions of readings from corresponding times in different *Ss*. It seemed wise, therefore, to apply some transformation to the data. A logarithmic rather than a percentage transformation was used since it provided a basis of comparison for initial levels as well as for changes. Each reading was, therefore, converted to a figure equal to its common log $\times$ 100 before any averaging or other statistical treatment was begun. The resulting distributions seemed by inspection to approximate the normal.

Analysis for stimulus-trace effects.—The important data of this experiment are the action-potential trends while the two stimuli of a pair are being presented and in the interval between them. Differences in these trends will reveal the effects of the several experimental conditione.

The first step in the analysis was to computs the mean reading for each second in the various groups and subgroups. The chance error of these measures was then determined by a series of analyses of variance. It was known from the graphic plots that beginning with the onset of the second tone the mean activity level rose to a much higher order in all groups. It seemed wise, therefore, to confine the analyses of variance to the readings before the onset of the second stimulus. As the graphic study (see Fig. 2) had already shown quite different trends in arm and jaw muscles, two separate analyses were made for each situation. A table was constructed in which each cell represented the mean reading for an *S* at a particular time and in a particular location; i.e., preferred or nonpreferred side and with a particular interval.

For Cond. 2 each cell was based on four readings; for the other conditions each was based on 16 readings. Besides individuals, the variables involved in each of these tables are time, electrode location, treatment (which hand was designated to perform a given role), and for Cond. 2, interval length. Since the problem was to evaluate the variances arising from time (i.e., the trends appearing in the graphs) the error variances used were the sum of all the interactions involving the time variable. (Certain of these time interactions might have been removed from the error variance; but were left in, giving perhaps somewhat greater generality.) Since the same individuals, locations, and treatments appear in the readings at each time, their systematic effects and interactions are of no concern for the evaluation of the differences between the

TABLE 1

TESTS FOR VARIATION OF ACTION POTENTIALS WITH TIME

Source	F	df	p	σ_M*
Cond. 1				
Responding arm	2.54	$\frac{7}{202}$	.05	1.5
Nonresponding arm	1.96	$\frac{7}{202}$		1.5
Masseters	2.64	$\frac{7}{217}$	0.5	1.3
Cond. 2				
Arms	12.5	$\frac{7}{1081}$	.01	.9** 1.0 .2 1.7
Masseters	13.9	$\frac{7}{1081}$	.01	.6 .7 .9 1.3
Cond. 3				
Arms	8.7	$\frac{7}{105}$	.01	1.4
Masseters	1.2	$\frac{7}{105}$		.9
Cond. 4				
Arms	25.8	$\frac{6}{92}$	.01	1.0

* Standard error of the mean for a given second.
** Respectively for Seconds 1–2, 3–4–5, 6–7, and 8.

time means. The F ratios and the standard errors of the time means resulting from these analyses are shown in Table 1. For Cond. 1, the time trends are evaluated separately for the responding and nonresponding arm, the variance between them and their time interactions having been removed from the error variance.

RESULTS AND DISCUSSION

Stimulus Trace Effects

The trends of the action potentials for the four conditions are shown in Fig. 2 in its several parts.

Trends associated with the second stimulus.—In all the curves there are large changes during and after the second stimulus. Where there is a choice reaction (Cond. 2, 3, and 4), there is an increase beginning with this stimulus and continuing evidently to the key pressing, which occurred at some time after the end of the graph. The arm used in pressing the key is compared with the arm not pressing in Table 2. There is an increase in both arms during and after the second stimulus. The jaw muscles act in similar fashion, though on a smaller scale. In contrast, where there is no choice reaction (Cond. 1), there is a difference in what happens in the two arms; while tension in the one is rising swiftly to an overt level, the tension of the other is falling. Here at least there is reciprocal action between these two muscle groups (extensors in the left and right arms), and the jaw muscles likewise become less tense as the active arm tension mounts. These trends are quite obvious, but some idea of their stability can be had from the variance analyses presented in Table 1. They would still be significant if the standard error of the means should increase by a large factor.

In the simple reaction situation a clear dominance of response in one muscle group appears by the end of the second stimulus; in the choice situation the response is spread through all muscle groups measured.

Effects associated with the first stim-

TABLE 2

COMPARISON OF ACTION POTENTIAL INCREASE IN ARM USED IN KEY PRESSING AND ARM NOT USED

Condition	First Stimulus*		Second Stimulus*	
	Arm Used	Arm Not Used	Arm Used	Arm Not Used
1	5.8	−1.2	30.3	−10.0
2	5.0	4.5	21.9	4.8
3	9.4	7.9	49.1	9.0
4	10.3	9.8	29.5	8.3

* Measures are in 100 log microvolt units and are differences between measures obtained after stimulus presentation and measures obtained before stimulus presentation.

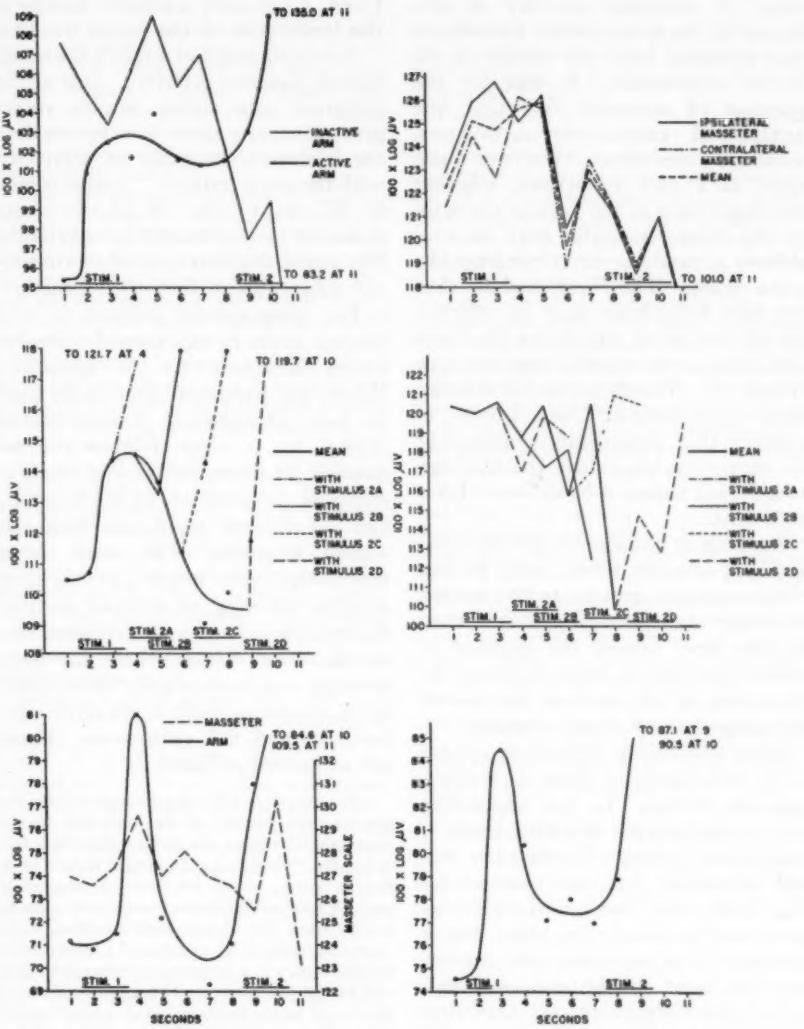


FIG. 2. *Upper left:* Mean action potential in the arms of 16 *S*s for 16 presentations of two stimuli with nonchoice response. The "active" arm is the one ultimately pressing key (Cond. 1). *Upper right:* Masseter responses corresponding to above (Cond. 1). *Middle left:* Mean action potentials in the arms of 32 *S*s with choice response and mixed intrastimulus-pair interval. For each *S* one-fourth of the trials were terminated at each of the intervals shown (Cond. 2). *Middle right:* Masseter activity corresponding to above (Cond. 2). *Lower left:* Masseter and arm activity for 8 *S*s, 16 trials each with 2-sec. stimuli and 5 sec. between stimuli of a pair (Cond. 3). *Lower right:* Arm activity for 8 *S*s, 16 trials each with 1-sec. first stimulus and 5 sec. between stimuli of a pair (Cond. 4).

ulus.—A muscular activity is produced by the first stimulus likewise, as was expected from the results of the earlier experiment. It was for the purpose of assessing this that the analyses of variance were carried out, as already described. We may question, as a null hypothesis, whether readings taken at the various times up to the second stimulus may be considered as random variations from the same value. The *F* ratios show that the null hypothesis may be rejected for all arm series, except for the inactive arm in the simple response case (Cond. 1). Trends in the jaw muscles show significance for Cond. 2 only. It appears that something, presumably the first stimulus, causes the readings at different times to be different from each other.

Looking at the curves for the arms in Fig. 2 one sees a rise, small for the choice-response groups, in the second following the onset of the stimulus. By the next second the increase is substantial and it seems to reach its maximum in all cases in the second following the end of the stimulus.

Some interesting differences appear in the subsidence of these excitations between stimuli. In the responding arm in the simple reaction there is little if any falling off before the second stimulus. For the choice situation where there was a variable intrapair interval (Cond. 2), there was a steady decline such that the tension was the same at the end of 5 sec. as at the beginning of the first stimulus. Where there were choice situations with a fixed intrapair interval (Cond. 3 and 4), there was a rapid decline after the peak, so that most of the intrapair interval was a period of stability. In the jaw muscles there is a slight rise of activity during the first stimulus, but its significance is questionable. The significant trend in Cond. 2 is mainly a general decline in the latter part of the period studied.

A striking incidental fact is the sidedness of masseter activity. Just as the preferred arm almost always shows greater activity than the other, so does the ipsilateral masseter as compared with the contralateral. This was true in 50 out of the 54 *Ss* on whom masseter measurements are available. The remaining four cases show consistent differences in the other direction.

The geographical pattern of the tension prior to the second stimulus varies according to the situation. Where the particular arm to be used in key pressing is predetermined (Cond. 1), it alone follows the sequence: (*a*) a tension rise with the first stimulus, (*b*) a decline in the interval, and (*c*) a large rapid rise with the second stimulus. The other arm, meanwhile, acts much like the jaw muscles, showing an over-all decline, if anything. In the choice situation, on the other hand, where the key-pressing arm is not rigidly determined by instructions, both arms show similar rise and fall sequences. These are compared in Table 2.

Characteristics of the stimulus trace.—From the general time pattern of the excitation in the responding members, one may conclude that this is not the response of a first-order system (19). Such a system would not show the regions of positive acceleration which these curves seem to display when they change their direction. The excitation could be the response of a second-order system, however, a response involving both first and second derivatives. The input to the system would be the auditory signal, a square wave of 1 or 2 sec. duration. The equations for the response of a second-order system would include two parameters; in one form of the equation they are ξ, the damping ratio, and ω_n, the natural angular frequency.

In the curves presented there are no evidences of oscillation. The damping ratio in the system must therefore be one or more, for this is an essential condition for nonoscillatory response. On the other hand, the damping cannot be as great as two, for if it were, the positive accelera-

tion would be so small it would not show as it does in the declining portion of the curve for Cond. 2. A rough estimate of the natural angular frequency necessary to produce the Cond. 2 curve with such an amount of damping would place it at something less than 3.5. The natural frequency would then be about .5 sec. and the period 2 sec.

With such a natural frequency, or anything like it, the maximum response in a critically damped or over-damped system would occur at the end of the signal. In all our cases the maximum is reached in the next second. This, however, would be expected if there is a delay due to transmission time. In the organism the time for transmission is large enough to take into account. Earlier studies (1, 3) showed the delay in the beginning of the response to be about .1 sec. and shifting the time scale by that amount would place the maximum in the second following the end of the signal.

We are confronted, however, not with one typical response curve, but with at least three varieties. In Cond. 1, there is a fast rise and slow decline; in Cond. 2 both are moderate; and in Cond. 3 and 4 there is a rather sharp though delayed peak. The shape of the response is modified according to the situation; the situation (instructions, previous responses, etc.) changes some parameter of the responding system and the stimulus then produces a response in this modified system.

But such a change as that from Cond. 2 to Cond. 1 could not be produced by any prior change of the two parameters of a second-order system. In Cond. 1 the rise is faster and the decay slower. Parameters which would speed up the rise would also make for faster decay. So there must be a shift of effective parameters during the course of the response. Such a shift would result from a feedback in the system, a feedback whose amount depended on the response. The existence of such a feedback somewhere in the system seems rather probable anyhow, and prior changes in its parameter would account for the major features of the response curve which change from situation to situation.

The traces studied here parallel the final responses in certain ways. They are located mainly in the same musculature, so far as our exploration goes. Where the two arms respond homonymously in the key pressing (Cond. 2, 3, and 4), they both show the rise and fall of excitation associated with the first stimulus. Where they act reciprocally in key pressing, they behave differently at the time of the first stimulus, although without clear reciprocity. Our stimulus trace itself is, of course, the same kind of action, so far as the records go, as the muscular contraction which results in key pressing. To the first stimulus there is, then, a response in many ways like the response given after the second stimulus and the recorded trace could as well be called an anticipatory response.

The excitation in the muscles has, therefore, certain of the characteristics postulated by Köhler (10), Hull (9), and Schlosberg and Heineman (18) for the stimulus trace. It is, however, subject to variation in time course and in geographical distribution according to the situation. In Köhler's type of analysis, these properties would probably be attributed to the stimulus trace if they affected the final judgments.

It is conceivable that the excitation here called a stimulus trace is later in the chain of events and more corrupted by extraneous factors than the processes hypothesized by Köhler and Hull. In some region of the cerebral cortex, for instance, there might be a process going on which answers to the exact description of the theoretical stimulus trace, but in the present state of knowledge this does not seem a necessary supposition.

It is still a question whether the muscular events play a determining role in the chain of happenings leading to the key pressing. This experiment gives no definite answer. From the results it can be seen that there are peripheral events which show the effects of the stimulation and of the "setting factors." Perhaps these are merely an overflow from the central nervous system important to the investigator as indices but of no consequence to the organism.

On the other hand, local conditions established in the muscle may prepare it to respond or not to respond and the proprioceptive pathway from the muscles could constitute the feedback pathway which seems to be necessary.

Relation of the Trace to Judgments

With so much resemblance between the action-potential results and the trace Köhler inferred, it seemed that his theory might be essentially correct, and that the muscular activity has the same characteristics as he referred to a brain state. But unfortunately for this conclusion, the judgments given in the present experiment do not follow the expected time function. The judgment percentages are shown in Table 3. (In view of the small number of judgments from each *S* and because of the need for averaging action-potential data, it was thought best to treat each group rather than each individual as a unit.)

TABLE 3

PROPORTION OF JUDGMENTS "WEAKER"

	% W Judgments	Total	*Ss*
Cond. 2			
.1-Sec. Interval	48	128	32
1-Sec. Interval	67	128	32
3-Sec. Interval	69	128	32
5-Sec. Interval	59	128	32
Cond. 3	65	128	8
Cond. 4	34	128	8

The percentages in this experiment cannot be converted into conventional time-error stimulus values. (They are not in a good deal of the standard literature of the subject.) We have available only one point on each of the psychometric curves, since only one stimulus value was used, and of course one cannot compute two parameters indicating the central tendencies and dispersions. From a mathematical viewpoint it would be possible that the various percentages are on psychometric curves all having the same intercept (time-error) but enormously different slopes; standard experiments, however, show that it is the intercept rather than the slope parameter which varies under such conditions as these. In any case it may be as well to study the responses given, rather than some mathematical abstraction from them.

We evidently have an aberrant series of percentages such as Harris (7) noted, and a stimulus trace inferred from them would bear very little relation to our action-potential data. It would be represented by a curve like that for the W judgments in Fig. 6. This outcome first appeared in data from 16 *Ss* in Cond. 2. On the suspicion that the percentages were unusual because of random error, the size of the group was doubled, but the larger group showed no material change over the smaller in either judgments or action potentials. Neither level of the first excitation, nor its rate of change, nor its ratio to the second bears much relation to the shift in judgment proportions from one interval to another.

At this impasse eventually a new hypothesis came to light. The arm muscles whose activity was being recorded might be looked upon as the terminal points of two reaction systems, both more or less in a state of excitation throughout the experiment. The final action of one of those (in "S system") is called by the experimenter a "Stronger" judgment, of the other (the "W system") a "Weaker" judgment. The records may show the degree of this excitation, and the relative amount of it in the two systems may be the determiner of the choice made. At the time the second stimulus comes, possibly the more active system has the advantage and will, other things being equal, press the key.

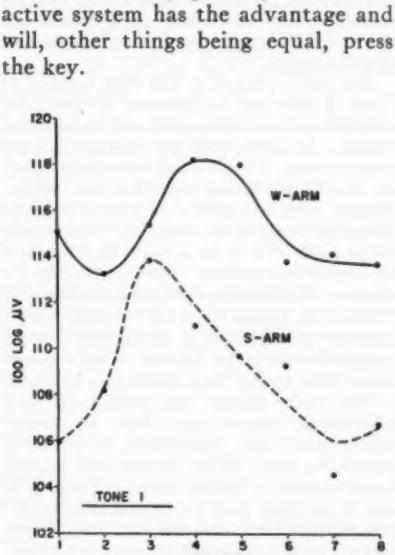


FIG. 3. Activity in arms designated for "Stronger" and "Weaker" judgments in Cond. 2 (variable intervals between stimuli). Based on data prior to the second stimulus only.

According to this notion the proportion of judgments should be a function of the difference between the excitations of the two systems, rather than of their absolute level. Exploring this possibility, a plot was made of the excitation in the S and W systems for Cond. 2 as a function of time (Fig. 3), with smoothed curves drawn by inspection, but with regard to the unequal precision of the points. The difference between the two does appear to vary with time: there is a difference in the shape of the S and W curves, so that at some points of time they are closer together, at others farther apart. At this point a statistical test of the difference in trend is needed.

This was secured by analysis of variance, and the procedure for performing these analyses was basically the same as that described previously. The question now, however, is whether the time effects (up to the beginning of the second stimulus) are different in the two systems. In technical language this is to ask whether the interaction between response systems and time is significant. To make the calculations a little less intricate, new tables were prepared in which each entry represented the difference between the S and W arms for one *S* for one point of time for presentations with one interval. Since only readings prior to the second stimulus were used, the table for Cond. 2 had unequal numbers in the columns representing the time variable, but this is no bar to the analysis of variance. With the tables consisting of W–S differences the question is whether the means of these for the various times are different. There were, it will be recalled, two treatments accorded *S*s; for half of them the right hand was designated S, the left W, for the other half the designations were reversed. The proper error term for our test, therefore, would be the pooled individual differences in time trends within these two treatments. The *F* ratio so computed for Cond. 2 is 2.47 (7 and 503 *df*) which has a level of significance between the 1% and 5% points. Briefly the difference in the trends is such that the difference (W–S) is at a minimum in the second following the stimulus, that it then increases to a maximum when the S curve reaches its starting level, then decreases and approaches a stable value as the W excitation dies out.

With the trend differences in the two systems ascertained, the next question was whether judgment proportions reflect them. In the data from Cond. 2 it is possible to make a comparison, because the sequences were terminated and judgments called for at a variety of times. (While the recording of action potentials is continuous, the proportion of associated judgments can, of course, be found only by ending a certain number of runs at a given time.) Figure 4 is a plot of the two variables, proportion of judgments and W–S system difference (in action potentials) each as a function of time. The action-potential differences are those read from the smoothed curves in Fig. 3. As there are but four points to compare, it was desirable to lose a minimum of degrees of freedom in plotting. The action-potential differences are therefore plotted on the a priori assumption that a zero difference would correspond to a 50% division of judgments. A second anchor point for the scale was obtained by putting the maxima of the two functions at the same point.[1] Although minor variations in the action-potential curve are, of course, dependent on the smoothing operation, there is evident a very good agreement between the functions. Turning

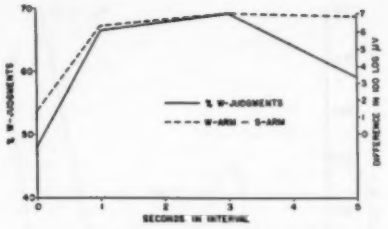


FIG. 4. Comparison of per cent W judgments with action-potential difference between the two arms just prior to the onset of the second stimulus (Cond. 2, variable stimulus)

[1] The maximum difference at point 5, as seen in Fig. 3, does not enter this comparison, since no judgments were taken at this point.

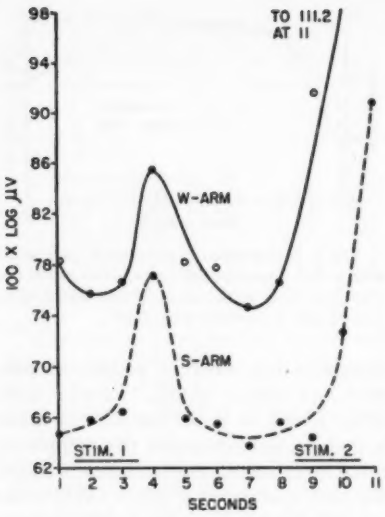


FIG. 5. Activity in arms designated for "Stronger" and "Weaker" judgments in Cond. 3 (constant 5-sec. interval and both stimulus durations 2 sec.)

to Cond. 3, it may be seen that the final W–S difference is in favor of the W system somewhat more than in Cond. 2. As would be predicted, the W judgments predominate (see Table 3).

In Fig. 5 the S and W curves for Cond. 3 show the same difference in trend as found in Cond. 2, though their statistical evaluation is not so reassuring. Computed in the same way as before, an *F* ratio for the difference in trend is only .8, indicating that other individuals in the same circumstances might not show the same difference at all. However, such a test takes no account of the sequence of differences, and perhaps should not be expected to show significance where the response is as brief as it is here. On the other hand, the minimum difference occurs where it would be expected to, and this is followed by a maximum about where it would be predicted. As would also be expected, the W excitation exceeds the S excitation just before the second stimulus. However this excess is brought about, it is followed by a preponderance of W judgments (65.9%).

The evidence so far is quite favorable to the hypothesis, but it lacks any case in which the action-potential difference is in favor of the S system. Actually it was at this point that the experiment called Cond. 4 was undertaken. The reasoning was that the S-system response (action potential) as seen in Cond. 2, and possibly in Cond. 3, seemed to have a more rapid rise and decay than the W system. But with a first stimulus lasting as long as 2 sec. the W-system activity has time to reach just as high a peak; its slower decay then puts it at an advantage during the subsequent 5 sec. It seemed possible, therefore, that terminating the first stimulus at 1 sec. would prevent the W system from reaching its peak and give the S

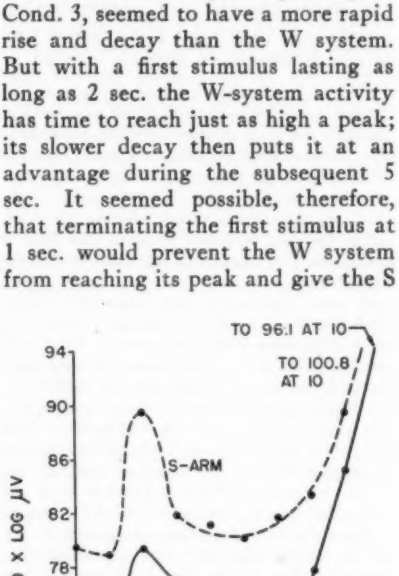


FIG. 6. Activity in arms designated for "Stronger" and "Weaker" judgments in Cond. 4 (constant 5-sec. interval with a 1-sec. first stimulus and a 2-sec. second stimulus)

system an advantage through the subsequent period. Under these circumtances, it was thought, there might also be the usual preponderance of S judgments (negative time error). (Experimenters reporting these have apparently used short-duration stimuli.)

As can be seen from the Cond. 4 curves in Fig. 6, the expectation about the shape of the S curves is borne out fairly well, though here again the *F* ratio is not at all significant. The difference follows the same course as in the two preceding situations with the W–S value at a minimum (algebraic) in the poststimulus reading, this being followed by a maximum, then by an approach to stability. The maximum and minimum here differ at the 5% level of confidence according to a *t* test making use of the general error term. As predicted, the S system is definitely more active than the W system at the beginning of the second stimulus. As Table 3 shows, there is also the expected majority of S judgments.

Neglecting the factor of time, the relation between system difference and judgment proportions may be studied by comparing all three groups. The difference in action potentials in the S and W arms just before the beginning of the second stimulus was used as the basis for plotting Fig. 7. (The four middle points, taken from Cond. 2, are of course derived from the same *Ss*; the end points are not.) It may be seen that the relation between system difference and judgment proportions extends over different experimental situations.

The over-all difference in level between the S- and W-system curves in Cond. 4 was not anticipated, though it might have been. In Cond. 2 and Cond. 3 the W system is at a higher level than the S system at all points, even before the onset of the standard stimulus; in Cond. 4 it is the S system

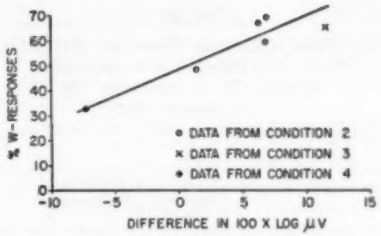


FIG. 7. Per cent W judgments as a function of the action potential difference in the two arms just prior to the second stimulus

which is always higher. It is of interest to find whether the differences between the S and W systems just prior to the second stimulus are significant. For this comparison the error term needs to include the variation among individuals within a treatment. In Cond. 2 individual differences were extremely large, with an *SD* of 36; the Cond. 3 experiment which was carried out later has an *SD* of 18; and the last experiment has an *SD* of only 6. This remarkable shrinkage may represent improvement in techniques on the part of *E*. (For testing the mean at one point, the *SD* between individuals is augmented a small amount by the error within *Ss*.) Though the mean W–S differences in the three situations are of the same order of magnitude, the *F* ratio for the Cond. 2 is less than 1, for Cond. 3 is 2.99 ($p = .07$), for Cond. 4 is 10.1 ($p = .01$). It seems possible that W–S differences before the second stimulus are typical of the population, but the Cond. 2 and Cond. 3 experiments are too imprecise to establish them.

Because of these differences in level it would be possible to predict the proportion of judgments the *Ss* will give even before the standard stimulus has been given. By the higher level of activity in one system than in the other the *Ss* show they already have a

TABLE 4

PER CENT RESPONSES GIVEN BY PREFERRED HAND AND MEAN ACTION-POTENTIAL DIFFERENCE IN FAVOR OF THE PREFERRED ARM

	% Judgments Given by Preferred Hand	Action-Potential Difference	N
Cond. 2			
.1-Sec. Interval	51.6	33.0	128
1-Sec. Interval	56.3	29.4	128
3-Sec. Interval	53.2	32.2	128
5-Sec. Interval	50.8	33.0	128
Cond. 3	58.5	12.7	128
Cond. 4	54.6	15.5	128

bias in favor of the one alternative. This preliminary bias seems quite puzzling at first, but the trend curves are, of course, averages taken from 16 trials, and it may be that the initial difference is one which develops during the series. If a response in one reaction system were to leave behind it an excitation lasting until the next pair of stimuli comes, a bias would tend to develop. Data in the present experiment are not adequate to test the notion, but Livingston (12) found such an effect in a similar situation. He reports that muscular activity in the arms during a choice reaction-time foreperiod is a function of the runs of stimuli (and responses) for some distance back in the series. It has also been shown that the proportion of judgment of pitch is similarly affected by prior runs (11). It seems as though a persistent excitation in one reaction system might be the immediate determinant of these nonrandom trends in judgment.

In the treatment of the data thus far, the difference between the preferred and nonpreferred arm activity has been removed from consideration by the averaging of responses from preferred and nonpreferred arms. The activity of the preferred arm is almost universally greater than that of the nonpreferred, up to the beginning of the second stimulus of each pair. To see whether this difference might lead to a majority of preferred hand responses, regardless of other factors, Table 4 was prepared. From this it does seem that *S* will tend to make the final response with the preferred hand, no matter whether it signifies "Weaker" or "Stronger" by instructions.

The relation of stimulus traces to judgments.—The experimental results support the following interpretations.

1. By instructions alternative response systems are established in *S*. The level of action potentials from the principal responding muscle in each is an indicator of the state of excitation of each.

2. The response system which makes use of the preferred arm will have a higher level of excitation than the one making use of the nonpreferred.

3. When an overt response has occurred in a given system, that system may be left in a state of higher excitation for at least a minute thereafter.

4. When the first stmulus of a pair is given, an excitation of both response systems occurs. This will reach its maximum within a second after the end of the stimulus and decline thereafter. (The exact course of the rise and fall depends on the situation set up by *E*.)

5. The systems designated "S" and "W" have different excitation trends. The "S" system seems to be more volatile; it has a quick rise and more rapid decay. The "S" system develops excitation more slowly and retains it longer. Early in the intrapair interval, therefore, with other things equal, the S system will be more active than the W system; later the opposite will be the case. (While this is undoubtedly true for our Cond. 2 with mixed intervals, and may be generally true for our Cond. 3 and 4 with fixed intervals, it probably is not true for all experimental situations.)

6. The second stimulus revives the excitation in both systems and raises it to a very high peak following its cessation. In one of the systems, in fact, the excitation is so great that the muscle develops enough force to close the response key.

7. The system which is most excited by the second stimulus is the one which is already most active. Hence, on the average, the number of "judgments" made one way or the other (stimuli being equal) will depend on the difference between the excitation levels in the two systems

at the time the second stimulus is given. The judgments will therefore be related to those factors which in turn determine the relative excitation levels of the two systems.

This determination of final response by prior level of system activity resembles the finding of an earlier study (3) in which the muscular reflex response to an auditory stimulus was shown to be related to prior tension level. Somewhat similar is the result of Bosma and Gellhorn (1) that the muscular stretch produced by loading a muscle determines a greater contraction resulting from cortical stimulation. The greater effect on the active muscle in the present case is more than proportional to its absolute activity level, for the procedure of expressing the results in logarithmic units would allow for such proportionality. In a study of response at the level of the absolute threshold level (5) it was proposed that the activity of a system, once it reaches a certain point, initiates a positive feedback which builds up the excitation to an overtly effective contraction. (A reduction of negative feedback would, of course, do as well.) Such a variable feedback was also suggested to account for the variations in trace excitation associated with the first stimulus of the present experiment. If the proportion of feedback increases with amount of excitation, the activity of a system would tend, after a certain point, to build up to whatever limiting point exists.

This interpretation is similar to the usual stimulus-trace theory of the time error in that it makes the proportion of judgments dependent upon the fading effects of the first stimulus. It differs radically from that theory in that the excitations considered are not in cortical fields exclusively, and in that the difference between two "traces," both started by the first stimulus, is considered to be the determiner, rather than the intensity of one trace. (In either case, of course, the intensity of the excitation produced by the second stimulus would have to play a role whenever it is allowed to vary substantially.) The notion of "competing" traces is favored because the single trace, if it is identified with the observed excitations, fails to account for the data. The response-system idea is favored in the first place because pertinent effects were observed in the terminals of such systems. A second advantage of the theory is that it leads to an explanation of the key pressings which are observed, rather than to an explanation of a state of awareness which is presumed to be indicated by these key pressings. It does not require that a cortical condition should be exactly reflected in the overt actions.

The difference in the way the S and W systems respond to the first stimulus raises further questions. The difference is evidently not anatomical, since each arm can be made a part of either system. Possibly the difference is determined by the same linguistic habit which makes the expression "stronger or weaker" more usual than "weaker or stronger."

There are certainly other factors taking part in response determination besides the ones here studied. Probably some operate by affecting the prestimulus level of various response systems, some by affecting the excitation aroused by the first stimulus, some by affecting the trace decay, or by adding new excitations somewhere in the decay period. Conditions coincident with the second stimulus (e.g., its intensity) or posterior to it, would presumably act by changing its effects. But, in general, in the theory presented anything affecting the response proportions and operating ahead of the second stimulus would have to do so by modifying the activity of one or more of the many competing response systems in the individual. This might be true also in those other situations which consist of a sequence of stimuli as well as in the intensity-judgment case.

Summary

Four groups of *Ss* were presented with pairs of tones under four different experimental procedures with *Ss* pressing a key in response to each pair in certain of the conditions constituting a judgment of the relative interests of the second stimulus. Search for stimulus-trace effects was made in the concurrent muscular activity in four body locations by electromyographic methods, and the relation of some of these to ensuing judgments was investigated.

Under all the experimental conditions there was an increment in muscular activity associated with the first stimulus of each pair, at least in those muscles likely to be involved in the subsequent key pressing.

For the simple reaction situation the excitation was highly localized; for the choice-response situations it was more widespread.

The temporal course of the excitation is somewhat like that usually attributed to the stimulus trace. It varies with the following factors: (*a*) simple vs. choice response; (*b*) regu-

larity vs. irregularity of time interval between stimuli of a pair; and (*c*) duration of the first stimulus of a pair.

It is, therefore, less closely bound to the stimulus than the stimulus trace is sometimes supposed to be, but varies with the "setting factors" as well. This variation could be accomplished by a feedback in the system.

The proportion of judgments of "stronger" and "weaker" bears no discernible relation to its total amount of activity in the two arms prior to the judgment.

It is related, however, in a variety of circumstances, to the difference in prior activity of the arms designated to give "stronger" and "weaker" judgments, the more active arms giving the greater number of responses.

It is proposed that the action potentials from the two arms represent the activity state of two competing response systems; that both are excited by the first stimulus but have differing response curves and initial levels; that, other things being equal, the system which has advantage at the time of the second stimulus will further build up its activity until enough force is generated to close a key.

(Received April 5, 1952)

References

1. Bosma, J. F., & Gellhorn, E. Electromyographic studies of muscular coordination on stimulation of the motor cortex. *J. Neurophysiol.*, 1946, **9**, 263–274.
2. Davis, R. C. An integrator and accessory apparatus for recording action potentials. *Amer. J. Psychol.* 1948, **61**, 100–104.
3. Davis, R. C. Motor effects of strong auditory stimuli. *J. exp. Psychol.*, 1948, **38**, 257–275.
4. Davis, R. C. Responses to 'meaningful' and 'meaningless' sounds. *J. exp. Psychol.*, 1948, **38**, 744–756.
5. Davis, R. C. Motor responses to auditory stimuli above and below threshold. *J. exp. Psychol.*, 1950, **40**, 107–120.
6. Freeman, G. L., & Sharp, L. H. Muscular action potentials and the time-error function in lifted weight judgments. *J. exp. Psychol.*, 1941, **29**, 23–26.
7. Harris, J. D. The effect of interstimulus interval on intensity discrimination for white noise. *Amer. J. Psychol.*, 1949, **62**, 202–214.
8. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
9. Hull, C. L. Behavior postulates and corollaries—1949. *Psychol. Rev.*, 1950, **57**, 173–180.
10. Köhler, W. Zur Theorie des Sukzessivvergleichs und der Zeitfehler. *Psychol. Forsch.*, 1923, **4**, 115–175.
11. Koester, T., & Schoenfeld, W. The effect of context upon judgments of pitch differences. *J. exp. Psychol.*, 1946, **36**, 417–430.
12. Livingston, W. A. Action potential measurements from the arms in the foreperiod of reaction time to visual stimuli. Unpublished Ph.D. dissertation, Indiana Univer., 1949.
13. Payne, B., & Davis, R. C. The role of muscular tension in the comparison of lifted weights. *J. exp. Psychol.*, 1940, **27**, 227–242.
14. Peak, H. The time order error in successive judgments and in reflexes. III. *Psychol. Rev.*, 1940, **47**, 1–20.
15. Postman, L. The time-error in auditory perception. *Amer. J. Psychol.*, 1946, **59**, 193–219.
16. Postman, L., & Page, R., Retroactive inhibition and psychophysical judgment. *Amer. J. Psychol.*, 1947, **60**, 367–377.
17. Roush, R. G. Tone burst generator. *Electronics*, 1947, **20**, 92–96.
18. Schlosberg, H., & Heineman, C. The relationship between two measures of response strength. *J. exp. Psychol.*, 1950, **40**, 235–247.
19. Trimmer, J. D. *Response of physical systems.* New York: Wiley, 1950.
20. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.

THE PROJECTIVE MEASUREMENT OF EXPERIMENTALLY INDUCED LEVELS OF SEXUAL MOTIVATION[1]

RUSSELL A. CLARK

Wesleyan University

The primary purpose of the present research has been to investigate the measurement of sexual motivation by means of the TAT. The impetus for this study came from some findings by McClelland *et al.* All of the TAT scoring categories employed by McClelland and his associates (3) for measuring the need-for-achievement (*n Achievement*) varied as a direct function of increase in motivation. That is to say, with an increase in experimentally induced strength-of-achievement motivation there was a corresponding increase in the frequency of appearance of the various achievement-related categories. This seems to indicate that for *n Achievement* there is no such thing as repression, suppression, or any sort of inhibition operating. This finding seems fairly plausible in view of the fact that in the American culture individuals are widely encouraged to work hard, get good grades in school, advance themselves in the business world, etc. There are few taboos attached to the attainment of achievement goals. However, these findings involved mean differences in frequencies based on groups of *S*s. A further study by McClelland *et al.* (3) involving case histories and psychiatric interviews indicated that there were a few individuals who were highly motivated to achieve but so anxious about achievement that they inhibited the expression of manifest achievement imagery in their TAT stories. This tentative finding ran contrary to the common assumption that projective measures circumvent the usual defenses of the individual.

For investigating the degree to which inhibition occurs in the TAT, it seemed desirable to select some motive which was apt to differ from *n Achievement* with regard to amount of anxiety that was apt to be involved. The sex motive, in our society at least, seems to have the necessary qualifications. There are all sorts of parental, social, and religious prohibitions against sexual activity. Because of this basic difference, and because it seemed possible to manipulate sex motivation experimentally, this motive was chosen as suitable for investigation.

PROCEDURE

Three separate experiments (A, B, and C) were conducted in the course of the present investigation. Each experiment involved two groups of male college *S*s one of which (experimental) took a group TAT after having been exposed to some type of sexually arousing stimulus. The other group (control) consisted of comparable *S*s who took the group TAT without prior sexual stimulation.

The TAT pictures employed were from the Murray 1942 series with one exception. They were presented in the following order: 7BM, 12BG, 14, 20, 10, 6F, Picture of a man working late at night in an office, and 6BM. Not all of the groups saw all eight pictures. The first five were shown for Exp. A, the first six for Exp. B, and all eight for Exp. C. The instructions which were employed for the administration of the group TAT are the standard instructions used by McClelland *et al.* (3, 4).

Experiment A.—The experimental group was

[1] The present paper is a somewhat different form of a dissertation presented to the faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy. The author is grateful to Drs. C. I. Hovland, I. L. Child, I. L. Janis, and D. C. McClelland for assistance in the planning and execution of this project.

first exposed to a series of photographic slides of attractive nude females and immediately thereafter given the TAT. The control group took the TAT immediately after having been exposed to a series of slides of landscape scenes, architecture, fashionably decorated rooms, etc. All slides used in this experiment were black and white.

The *S*s were in two sections of an elementary psychology class. They were tested before there had been any formal discussion of projective techniques. Two *E*s (graduate students) met the control class; one to conduct an "investigation on factors affecting aesthetic judgment," which was the rationale for the control slides, and the other to conduct "a standardization of a test of creative imagination." The *S*s were asked to rate each slide with respect to degree of attractiveness stating what it was about the picture that makes it attractive or unattractive. They used mimeographed line rating scales with spaces for comment. All slides were presented briefly in order to enable *S*s to get an over-all impression, and then on the second presentation the slides were shown for a longer period thus enabling *S*s to make the ratings and pertinent comments. The TAT was given by the second *E* immediately thereafter. Every effort was made to convince *S*s that the rating of the slides and the TAT were separate experiments.

The experimental group was tested on the next day by the author and one of the previous *E*s. After again doing a variety of things to give *S*s the impression that the presentation of the slides and the TAT were two different experiments, the author gave the rationale for the presentation of the nude slides, which was an attempt to make the presentation of the nude slides seem very plausible to the *S*s. This rationale referenced the work of Sheldon (6) on the correlation between body type and certain physical and mental disorders, explained why the pictures of nude females were drawn from art photography (because of social pressure against standardized body-typing photographs), and requested *S*s to use a mimeographed line rating scale with space for comments in evaluating the eight pictures to be shown. The *E* indicated that comments such as "breasts just the right size and shape" and "legs long and shapely" were adequate. In order to further set *S*s for sexual stimulation by the slides, *E* then said:

> I think that you will find that most of these pictures are of at least fairly attractive girls. We have purposefully included mostly those body types that we feel should be fairly appealing. So when you use the rating scale, don't be concerned if many of the ratings fall towards the attractive end of the scale, but try to make fine discriminations among the more attractive girls. Now remember your job is to judge the sexual attractiveness of these girls. I don't think you will find this an especially unpleasant task.

The nudes were projected life-size on a large screen. After *S*s had finished their ratings and pertinent comments about these slides, *E* collected the papers, packed up the slides, and left. The other *E* then came forward and administered the TAT with the same remarks and instructions that were given the control group.

Experiment B.—Basically this experiment consisted simply of employing a very attractive female as administrator of the TAT for the experimental condition and a male administrator for the control group. No nude or neutral slides were involved. It was designed to determine whether the presence of an attractive, clothed female would elicit a type of sexual phantasy less apt to be guilt-ridden than that obtained in Exp. A.

The *S*s were members of the freshman class who, for one reason or another, had missed taking the TAT as a part of a battery of personality and aptitude tests. The *S*s for the experimental and control groups were assigned at random and received notices from the Dean's office to report for the test. These *S*s were run in small groups of 4, 5, or 6 over a four-day period, with the male *E* (control) testing on the first two days, and the female *E* (experimental) testing on the last two days. All testing was at night and in an office, in order to assist in making the female's presence less formal and more intimate. The girl was dressed attractively and wore an allegedly stimulating brand of perfume. She had trouble with the projector on the first TAT picture in order to enable her to request the males to come to her assistance. They gathered around and did so. The privacy of *S*s while writing was insured by *E*, after having projected the TAT slide for 20 sec., retiring to a remote easy chair during the 5 min. allowed for writing about each TAT slide. Also, each *S* deposited his protocol in an envelope upon completion of the test.

Experiment C.—It is a very common assumption that alcohol reduces guilt, fear, or anxiety. In fact, recent work by Conger (1) with white rats seems to demonstrate that this is the case. Experiment C employed alcohol in an attempt to reduce sex-involved guilt or anxiety. The experimental group took the TAT in a fraternity beer-party setting after having been exposed to nude slides as compared with a control group also in a

beer-party setting but without exposure to any kind of slides.[2]

Two presumably comparable fraternity groups were used as the experimental and control groups. The test was rationalized on the ground that it was desirable to determine the effects of informal environment and alcohol on the test of "creative imagination." The test was given approximately 1½ hr. after the start of the beer parties, and *S*s were permitted to continue drinking during the test. The nude pictures were exposed on slides to the experimental group through clandestine arrangements designed to dissociate these pictures from the later TAT. The TAT was given shortly after the last of eight nude pictures was shown.

Scoring.—The scoring was done by the author and one other person. The author rescored the stories several times. The other scorer practiced on 150 stories under the author's tutelage before beginning the reliability check. In the case of discrepancies the final scoring was decided upon after a joint conference of the two scorers.

The first analysis of the TAT protocols was to reveal manifest sex imagery. Sexual relationships were categorized as (*a*) primary—explicit or implicit evidence for sexual intercourse, (*b*) secondary—evidence for the occurrence of such secondary sex activity as kissing, dancing, fondling, etc., (*c*) tertiary—characters in the stories perceived as sweethearts, on a date, courting, in love, etc., but not engaged in either primary or secondary sexual activity. Sexual activity was scored only once for each story, and that which was biologically most sexual received priority.

In addition to sex imagery, the protocols were also scored for guilt. A decision was made to score for guilt only in connection with sexual activity. It could be argued that sexual guilt might easily appear in a disguised form by being displaced to other types of activity. The author made an initial attempt to score for guilt regardless of the activity involved and obtained no group differences whatsoever. Three general categories are scored as being indicative of the presence of guilt: (*a*) someone is ashamed, guilty, sorry, anxiety-ridden, morally concerned, embarrassed, etc. over sexual activity; (*b*) someone is punished, criticized, ostracized for sexual activity; and (*c*) someone punishes himself in some concrete fashion as a result of a sexual activity. Guilt is also scored only once for each story.[3]

The per cent agreements between the two scorers were 91% for sexual imagery and 90% for guilt over sex. These percentages were obtained by considering agreements to be those categories which both scorers had marked as being present. Those categories which one scorer had marked as being present but the other scorer had marked as absent were counted as disagreements. Those categories which both scorers had indicated as being absent were not considered in the above calculations. Thus, agreements of the present-present variety are considered but not agreements of the absent-absent variety.

Results

The chi squares in this study were calculated in the following manner: The frequency of appearance of any particular category was calculated for each *S*. The two groups under comparison were combined, and the dividing point for setting up the contingency tables was determined by selecting that frequency which most nearly enabled the placing of 50% of all *S*s above and below that dividing point. Yates's correction for discontinuity was used in every case regardless of the magnitude of the cell entries.[4]

Table 1 demonstrates that the experimental groups for Exp. A and B exhibited significantly less manifest sex imagery than the controls. The experimental group for Exp. C, how-

[2] Neutral slides were omitted from this control group because a few stories in Exp. A seemed to be determined by a set for evaluation of the aesthetic qualities of the TAT slides which presumably was established by having rated the aesthetic qualities of the "control" slides.

[3] The stories were also scored for *affect* over sexual activity. A three-way breakdown was made involving: (*a*) passion, lust, etc.; (*b*) love, tenderness, etc.; and (*c*) no affect present. The results of this analysis are the same as those that obtain for the sexual activity and are not reported because they add little to the interpretation.

[4] There were five TAT pictures used in Exp. A, 6 pictures in Exp. B, and 8 pictures in Exp. C. In each case calculations are based on the maximum number of pictures common to the groups being compared. Also, in examining the tables that follow, the reader should be warned that the per cent of *S*s above the dividing point for a given group may vary from one comparison to another because a different dividing point is often involved.

TABLE 1

PER CENT OF *Ss* ABOVE THE DIVIDING POINT IN RESPECT TO THE FREQUENCY OF ALL SEX IMAGERY

Condition	*Ss*	% *Ss* Above Dividing Point	x^2
Exp. A			
Exp.	40	27.5	10.12**
Control	38	65.8	
Exp. B			
Exp.	30	30.0	7.34**
Control	29	69.0	
Exp. C			
Exp.	35	60.0	5.85*
Control	27	25.9	

* Significant at 5% level.
** Significant at 1% level.

ever, showed reliably more sex imagery than the control group.

Table 2 presents the comparisons of Exp. A and Exp. C. The reader is reminded that theoretically these two experiments differed only with respect to the fact that one of the conditions involved a beer-party atmosphere. In practice, however, there were certain minor differences other than this: The *Ss* for Exp. A were mainly sophomores, with a few juniors and seniors. The *Ss* for Exp. C represented, of course, a cross section of the four classes. Also, the nude slides were achromatic for Exp. A but chromatic for Exp. C. Finally, the control group for Exp. A was exposed to neutral slides, while the control group for Exp. C was not exposed to neutral slides. Interexperiment comparisons involving Exp. B are not presented mainly because the sexual-arousal stimulus differed markedly from that employed for the other two experiments, and only freshmen were employed as *Ss* in Exp. B.

Table 2 indicates that in general the groups of Exp. C (alcohol) show more sexual imagery than the groups of Exp. A (nonalcohol). Comparison of groups was made also with respect to the three types of sexual imagery—primary, secondary, and tertiary. The general finding was that primary sexual imagery was the main contributor to the differences in over-all sex imagery contained in Tables 1 and 2. Also, Table 2 shows that there is no significant difference in over-all sex imagery between the control groups of Exp. A and Exp. C. However, the control group of Exp. C does have a significantly greater amount of primary sexual imagery than the control group of Exp. A.

Tables 3 and 4 reveal that those groups which show more sex imagery (see Tables 1 and 2) also exhibit more guilt over sexual activity.

Table 5 presents the comparisons of the amount of guilt involved in the

TABLE 2

PER CENT OF *Ss* ABOVE THE DIVIDING POINT IN RESPECT TO THE FREQUENCY OF ALL SEX IMAGERY

Experiment A			Experiment C			x^2
Condition	*Ss*	% *Ss* above Dividing Point	Condition	*Ss*	% *Ss* above Dividing Point	
Exp.	40	27.5	Exp.	35	77.1	16.60**
Control	38	15.8	Exp.	35	45.7	6.51*
Exp.	40	27.5	Control	27	63.0	6.89**
Control	38	65.8	Control	27	63.0	

* Significant at the 5% level.
** Significant at the 1% level.

TABLE 3

PER CENT OF *Ss* ABOVE THE DIVIDING POINT IN RESPECT TO THE FREQUENCY OF SEX-INVOLVED GUILT

Condition	*Ss*	% *Ss* Above the Dividing Point	x^2
Exp. A			
Exp.	40	20.0	5.31*
Control	38	47.4	
Exp. B			
Exp.	30	16.7	8.01*
Control	29	55.2	
Exp. C			
Exp.	35	51.4	5.69*
Control	27	18.5	

* Significant at the 5% level.

different types of sexual imagery. It is clear that more guilt is evoked by primary sex than by secondary or tertiary sex. One also might expect that stories involving secondary sex would contain more guilt than those involving tertiary sex. This, however, is not the case. A number of explanations could be offered for this inconsistency. For example, one of the TAT pictures employed is usually perceived as a man and woman in an embrace. This picture elicited a large number of stories concerning a husband and wife in an embrace or kissing. This type of story would, of course, be scored as secondary sex. In our society, for husband and wife there are no moral sanctions against this type of activity. Also, some of these same stories involved embracing (secondary sex) but this was in a context of consolation for the loss of a son in the war, etc. These stories were scored in order to be consistent with the criteria, but in future work it

TABLE 4

PER CENT OF *Ss* ABOVE THE DIVIDING POINT IN RESPECT TO THE FREQUENCY OF SEX-INVOLVED GUILT

Experiment A			Experiment C			x^2
Condition	*Ss*	% *Ss* above Dividing Point	Condition	*Ss*	% *Ss* above Dividing Point	
Exp.	40	20.0	Exp.	35	82.9	26.89*
Control	38	47.4	Exp.	35	82.9	8.62*
Exp.	40	20.0	Control	27	59.3	9.07*
Control	38	47.4	Control	27	59.3	

* Significant at the 1% level.

TABLE 5

PER CENT OF GUILT INVOLVED IN EACH OF THE THREE TYPES OF SEX IMAGERY FOR ALL OF THE GROUPS COMBINED

	Primary Sex Imagery	Secondary Sex Imagery	Tertiary Sex Imagery
Per cent of stories containing guilt	70.4	22.1	30.4
Standard error of per cent	3.0	5.0	4.1
	Primary Sex vs. Secondary Sex CR = 8.3 $p<.01$	Primary Sex vs. Tertiary Sex CR = 7.9 $p<.01$	Tertiary Sex vs. Secondary Sex CR = 1.2

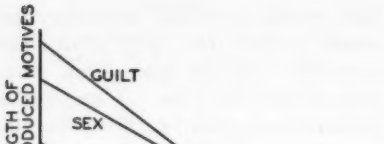


FIG. 1. Hypothetical curves showing the strength of sex and guilt motives evoked by the arousal conditions for the experimental and control groups

would be well to eliminate this type of story either by special provisions in the scoring system or by avoiding the use of this type of picture (TAT 10).

DISCUSSION

The two main factors considered in the present study are sex and sex-involved guilt. One logical way of interpreting these results involves an approach-avoidance conflict type analysis such as that proposed by Lewin (2) and Miller (5). However, before considering the results in the light of such an analysis, it is necessary to make a clear-cut distinction between two different aspects of both sex and guilt that are presumably involved in this study. First, there is the degree of sex arousal and accompanying guilt that is evoked by the experimental procedures involving the presence of an attractive female and the presentation of nude slides. These motives can be referred to as stimulus-produced sex and guilt motives. The second aspect concerns the sexual arousal and accompanying guilt that might result from the writing of stories of a sexual nature. These motives can be referred to as response-produced sex and guilt motives.

Figure 1 depicts the relative strengths of the stimulus-produced motives for the experimental and control groups. In this figure the guilt gradient is shown as being reduced by the consumption of alcohol. The control groups are assumed to be located toward the weak end of the stimulus continuum. The experimental groups are located at a point in the strong end of the stimulus continuum. This figure also assumes that the guilt or avoidance gradient is steeper than the sex or approach gradient. In other situations it has been found that avoidance gradients are steeper than approach gradients, and Miller (5) has discussed possible reasons for this. In the discussion that follows, the conclusions drawn depend, of course, on the shapes of the gradients and the location of the experimental and control groups with respect to these gradients. For the purpose of simplicity the gradients are assumed to be linear in nature. The exact location of the two groups with respect to these gradients is, of course, not known. If the positions depicted in Fig. 1 were changed at all markedly, some of the conclu-

FIG. 2. Hypothetical curves showing the strength of sex and guilt motives evoked by the writing of stories containing the various kinds of sex imagery

sions to be drawn would no longer hold. Figure 2 presents the strengths of the response-produced sex and guilt that are evoked by writing stories containing the various kinds of sex imagery. The considerations that apply to the gradients in Fig. 1 also apply here.

The question that arises next concerns the relationship between the stimulus-produced sex and guilt and the writing of stories containing sex imagery. The first assumption made is (*a*) that the expression of sex in the TAT stories is a direct function of the algebraic summation of the stimulus-produced sex and guilt gradients. That is, the higher the sex gradient above the guilt gradient, the greater will be the tendency to express sex. The second assumption is (*b*) that the guilt appearing in the TAT stories is a function of expressing sex in the stories and not a function of the guilt that is evoked by the arousal stimuli. There is some evidence to indicate that the second assumption is correct. If this assumption were true, there should be no difference in percentages of guilt over a given kind of sex imagery expressed in the TAT protocols between the experimental and control groups. That is to say, both groups should experience the same degree of guilt for writing a given kind of sex story. If the level of guilt evoked by the arousal conditions also has a tendency to be expressed in the stories, this should lead to a higher percentage of expressed guilt for the experimental group because of the greater strength of guilt due to the arousal conditions for this group. There is, however, no evidence for this. The percentage of guilt over primary sex obtained by combining the three experimental groups is 73.6%, and that for the combined control groups is 67.3%. The critical ratio for this difference is 1.01 with a probability value of .31.

In the light of these assumptions and considerations the findings now will be analyzed. The first major result to be explained is the finding that for Exp. A and B the control groups express more sex and guilt in the TAT stories than their respective experimental groups. By consulting Fig. 1 it can be seen that for these experimental groups the stimulus-produced guilt is at a higher level than the stimulus-produced sex. Just the reverse is true for these two control groups. Thus, one would predict that the control groups would have a greater tendency to express sex. Having expressed a greater amount of sex and especially a greater amount of primary sex, it follows that the control group should express more guilt. This can be understood by examining Fig. 2 and recalling the assumption that expressed guilt is primarily a function of writing stories containing sexual imagery.

The second major finding is that for Exp. C the experimental group expressed more sex and guilt than did the control group. Again by examining Fig. 1 the explanation for this can be sought. With a sufficient lowering of the guilt gradient due to alcohol, it can be seen that there is a greater difference between the sex and guilt gradients for the experimental group than for the control group. The line of reasoning for this difference in expression of sex and guilt is the same as that applied to the previous finding.

The last major finding to be considered concerns the interexperiment comparisons involving Exp. A and C. In general, the alcohol groups of Exp. C express more sex and guilt than do the nonalcohol groups of Exp. A. By checking Fig. 1 the reader can satisfy himself that for any of these

particular comparisons the alcohol group in question should express more sex than the nonalcohol group in question. However, when one examines Fig. 2, it can be seen that the alcohol groups should have less guilt over expressing sex because of the lowered guilt gradient in this situation. This, of course, does not conform with the empirical findings. On the other hand, one might expect a higher total frequency of guilt for the alcohol groups because of the higher total frequency of both over-all sex and primary sex. Nevertheless, according to Fig. 2, one would expect a lower percentage of guilt over a given kind of sex imagery. However, this is not the case. The percentage of guilt over primary sex obtained by combining the experimental and control groups of Exp. C is 68.6%. The comparable percentage for Exp. A is 65.0%. This difference yields a critical ratio of .43 with a probability value of .67. From this it might be concluded that the theory is wrong in assuming that alcohol reduces guilt over expressing sex. There is, however, at least one possible alternative hypothesis that could be checked by further work. This involves the consideration that differences actually exist in the stories which are categorized, for example, as primary sex. It is the author's impression from rereading the primary sex stories for the alcohol and nonalcohol groups that the stories of the former groups are more sexual in nature. That is, the statements "She became pregnant" and "They had an impassioned sex orgy" are both scored as primary sex, but the latter type of imagery seems, to say the least, a more powerful statement. A more refined scoring system or the present system modified such that sex imagery could be scored more than once per story might demonstrate this impression to be true. If the subjects in the alcohol groups are expressing a more primary type of sex, the increased guilt that this would involve might compensate for the decrease in guilt due to alcohol. This would account for the finding that there is no difference in the percentages of guilt involved in primary sex for the two conditions.

Summary

This study involved the projective measurement of experimentally induced levels of sex motivation. Sexual motivation was aroused by the presentation of slides of nude females in two instances and in another instance by the presence of an attractive female. The TAT protocols were analyzed for manifest sex imagery and sex-involved guilt.

The results showed that under normal conditions the experimental groups expressed significantly less sex and guilt in the TAT stories than did the control groups. Under conditions of alcohol these results were just the reverse. That is, the experimental group showed significantly more sex and guilt than did the control group.

These results are interpreted by assuming that under normal conditions the guilt evoked by sexual arousal is sufficient to inhibit the expression of sex with a consequent lowering of guilt. Under alcohol, however, the guilt over sexual arousal is reduced enough to permit the expression of sex with a resulting increase in expressed guilt.

(Received April 16, 1952)

References

1. Conger, J. J. The effects of alcohol on conflict behavior in the albino rat. *Quart. J. Stud. Alcohol*, 1951, **12**, 1–29.

2. LEWIN, K. Environmental forces in child-behaviour and development. In C. Murchison (Ed.), *A handbook of child psychology*. Worcester: Clark Univer. Press, 1931.
3. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. (On file at Wesleyan Univer. library.)
4. MCCLELLAND, D. C., CLARK, R. A., ROBY, T., & ATKINSON, J. W. The projective expression of needs. IV. The effect of the need for achievement on thematic apperception. *J. exp. Psychol.*, 1949, **39**, 242–255.
5. MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.
6. SHELDON, W. H. *Varieties of delinquent youth: An introduction to constitutional psychiatry*. New York: Harper, 1949.

EFFECT OF INCENTIVES UPON REACTIONS TO PERIPHERAL STIMULI [1]

HARRY P. BAHRICK
Ohio Wesleyan University

PAUL M. FITTS AND ROBERT E. RANKIN
The Ohio State University

Most recent investigations of the relations between motivation and perception, such as those of Bruner and Goodman (3) and of Levine, Chein, and Murphy (6), have dealt with changes in the manner in which given objects are perceived as the result of changes in need states or attitudes. The present investigation deals with a different aspect of the problem, the effect of changes in motivation on the probability that a particular stimulus will be responded to, i.e., the relation of motivation to the selectivity of perceptual processes.

We never respond to all of the available stimuli in our surroundings, but only to selected aspects of the environment. Not only are perceptual processes highly selective, but the focus of this selectivity may shift from moment to moment. It has been recognized for a long time that such perceptual phenomena are related to the selectiveness of the learning process, yet few studies of learning have been concerned with the prediction of perceptual responses. This may be due to the fact that such responses are not directly observable, but become accessible to empirical investigation only in so far as they can be inferred from changes in performance.

Recently, Berlyne (2) has attempted to extend the postulates of Hull's learning theory to permit the prediction of selective perceptual responses. However, the functional relations between such responses and motivational conditions need to be explored empirically in order to gain a better understanding of learning and performance of complex tasks.

The present study was planned to test the hypothesis that an increase in incentives results in an increase in the selectiveness of perceptual behavior, provided that the organism is familiar with the stimulus situation and has stable expectations with respect to the elements of the situation which can lead to need reduction. According to this hypothesis, a person who is offered a strong incentive would be expected to exhibit highly selective perceptual response and would be unlikely to perceive irrelevant stimuli, i.e., stimuli which are less related than others to his expectations of reward. Under the same circumstances, a person offered a weak incentive would be more likely to perceive irrelevant stimuli. It should be noted that by relevant we usually mean relevant to the goals and incentives specified by *E*. It is assumed that any object that an organism responds to must have some relevance for the organism.

More specifically, it is hypothesized that perceptual selectivity in a psychomotor task is influenced by (*a*) the incentives, or rewards, offered by *E* for "good" performance, and/or the punishment proposed for "poor" performance; and (*b*) how *S* interprets the

[1] This research was supported in part by the United States Air Force under Contract No. AF 33 (038)–10528 monitored by the Perceptual and Motor Skills Research Laboratory, Human Resources Research Center.

relation of the different aspects of the task to the incentives, i.e., what factors he expects to be most closely related to success or failure.

It is further hypothesized that a continuously changing aspect of a complex task will tend (other things being equal) to be interpreted by *S* as more important to success than an infrequently and intermittently changing aspect. It is also believed that a stimulus object in the peripheral visual field will tend (other things being equal) to be interpreted as less important than an object in the central part of the field. However, our concept of "peripheral" stimuli also includes all aspects of the central task that *S* classifies as being relatively unimportant.

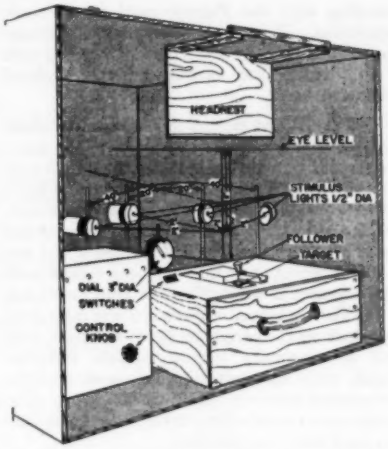


FIG. 1. The experimental apparatus showing the target and follower of the continuous tracking task, the peripheral stimulus lights and their switches, and the peripheral instrument dial with its control knob

METHOD

Apparatus.—A central continuous tracking task was used together with several intermittent peripheral tasks. A modified Buxton Division of Attention Test (4) provided the central task (see Fig. 1). The *S* was required to move a follower back and forth with his right hand so as to keep it in contact with a moving target. A timer at *E*'s control panel recorded the amount of the time that the follower remained in contact with the target. The *S* worked in a seated position, with his eye level 10 in. above the central tracking task.

Two kinds of peripheral stimuli were employed. One consisted of four lights, located in a semicircle in front of *S* at a distance of 24 in. away from and 8 in. below the eyepoint. One pair of lights was 20 degrees to either side of the median line of the body and the other was 40 degrees to either side. The lights were placed behind translucent windows, ½ in. in diameter, as shown in Fig. 1. The intensity of the light at the aperture was 24 ft.-candles and the contrast ratio between the aperture and its surroundings was 8:1.

Single lights were turned on from time to time by *E* by means of noiseless switches. The *S* could turn the lights off by operating the appropriate one of four switches. A timer at *E*'s control panel recorded the length of time each light remained on.

One of the instrument dials of the Grether Dual-Pursuit Apparatus (5) provided the second kind of peripheral stimulus. It was placed 24 in. away from and directly in front of *S*, and was 10 in. below eye level. The instrument pointer automatically moved off center six times during a period of 1 min. The extent of pointer deviation varied from 7 to 45 degrees. Both clockwise and counterclockwise movements occurred. The *S* could return the pointer to its neutral position by turning the control knob, as shown in Fig. 1, with his left hand. A timer at the control panel accumulated the time during each 1-min. trial that the pointer remained off center.

The entire apparatus was enclosed by a large booth which was open on the side facing *S*. A headrest was provided so that the eyes of all *S*s were maintained at a distance of 16 in. from the central task. The *E*'s control panel was behind *S*.

An *O* viewed *S* through a one-way screen in the front of the booth, and recorded the number of eye-fixation changes that occurred during the performance of some of the tasks. The *O* also recorded eyeblinks between trials, but the latter data are summarized in another report and will not be discussed in the present paper.

Subjects.—Eighty male and 20 female college students served as *S*s.

Experimental tasks.—The experiment required the performance of a peripheral task simultaneously with the performance of a central task by *S*s who were given one of two incentive conditions. The central task consisted in all cases of

tracking with the Buxton continuous pursuit apparatus. The score on this task was the number of seconds *S* remained on target during each 1-min. trial. Three variations in the peripheral task were used.

1. For the first task, *S* was not warned about the peripheral lights. At predetermined intervals of 15, 35, and 50 sec. during the third 1-min. tracking trial, three of the four peripheral lights were turned on, one at a time, for 5-sec. periods. At the end of this third trial, *S* was questioned as to whether he had noticed the lights, and was asked to report those that had been turned on in the sequence in which they had been presented.

2. For the second variation, *S* was instructed to watch for the peripheral lights and turn them out by operating the appropriate switches. Again three lights were turned on at the predetermined intervals during a 1-min. trial. The cumulative reaction time to the three lights during each trial was recorded.

3. For the third variation, *S*s were instructed to keep checking the instrument pointer of the Grether pursuit apparatus, and to adjust the pointer whenever they noticed that it was not on center. Six discrete pointer deviations occurred during each 1-min. test trial. The cumulative amount of time the pointer remained off the center position during each trial was recorded. This task differs from the preceding peripheral task in that it could not be accomplished through the use of peripheral vision. The *S* had to make an eye movement before he could detect a pointer deviation.

Procedure.—All *S*s received 13 1-min. trials followed by 30-sec. rest periods. Trials 1, 2, 4, and 7 were practice trials on which no special incentives were offered. On Trials 3, 5, 6, 8, and 9, *S*s in Group I were offered a bonus for good performance; on Trials 10, 11, 12, and 13, *S*s in Group II were offered a similar bonus.

All *S*s were paid a standard wage of 75 cents an hour. On bonus trials they were given a chance to increase their pay by earning high scores. Bonuses ranging from 5 cents to 3 dollars were offered on each high-incentive trial. The typical *S* earned a total bonus of 25 cents.

At the end of the second practice trial *S*s were assigned to one of the two groups in such a way as to equalize group means and *SD*'s for early performance on the central tracking task. This procedure, it should be noted, would not be expected to equate the groups for performance on the peripheral tasks.

At the start of the experiment each *S* was seated so that he could rest his head comfortably against the forehead rest, and was shown how to operate the apparatus. Instructions were given before each trial.[2] The special conditions for each of these trials were as follows:

Trials 1–2. All *S*s practiced the continuous tracking task under uniform conditions.

Trial 3. Group II was told that practice trials would continue. Group I was offered a money bonus for good tracking performance. Without previous warning three peripheral lights were presented while *S*s performed the central tracking task. At the end of the trial all *S*s were asked whether they had seen any lights, and if so, what the sequence was in which they had been turned on.

Trial 4. All *S*s practiced the continuous tracking task, and tried to keep the peripheral lights turned off by operating the switches. No bonus was offered either group.

Trials 5–6. Same as Trial 4, but Group I was offered a money bonus for good performance. The instructions stated that the time required to turn out the lights would be subtracted from the time-on-target tracking score in the computation of the total score.

Trial 7. All *S*s practiced the tracking task, and tried to correct for intermittent deviations of the dial pointer by adjusting a knob. No bonus was offered either group.

Trials 8–9. Same as Trial 7, but Group I was offered a money bonus. The instructions stated that the time required for centering the pointer would be subtracted from the total time-on-target.

Trials 10–11. Same as Trials 5–6, but Group II was offered a bonus instead of Group I. This was the first time *S*s in Group II knew that bonuses were being offered. Group I was given practice instructions.

Trials 12–13. Same as Trials 8–9, but Group II was offered the bonus instead of Group I.

Results and Discussion

The means and *SD*'s of scores made by the two groups for all trials on the central and peripheral tasks are given in Table 1. The significance of the differences between the mean scores of the two groups on each trial is indicated by *t* ratios.

The two groups had approximately equal performance scores on the central tracking task during the first two practice trials. However, on all of the subsequent trials on which they

[2] The complete instructions are available in (1).

TABLE 1

MEAN SCORES PER TRIAL FOR DIFFERENT TASK AND INCENTIVE CONDITIONS

$n = 50$ in each group

Trial	Peripheral Task	Incentive Condition		Central Task (sec. on target)					Peripheral Task[a]				
		Group I	Group II	Group I		Group II		t[b]	Group I		Group II		t
				Mean	*SD*	Mean	*SD*		Mean	*SD*	Mean	*SD*	
1	None	Low	Low	37.5	7.1	37.2	7.8						
2	None	Low	Low	43.4	6.7	42.9	6.9						
3	3 Lights (Verbal R.)	High	Low	48.5	5.4	43.4	6.3	4.4**	2.12	1.8	3.16	1.5	2.7*
4	3 Lights (Manual R.)	Low	Low	41.4	5.5	40.1	5.9	1.1	7.20	5.7	5.50	4.4	1.7
5–6	3 Lights (Manual R.)	High	Low	45.3	4.6	41.9	5.7	3.4**	5.35	3.2	4.87	3.0	.8
7	6 Pointer Adjustments	Low	Low	39.1	5.5	36.5	5.6	2.3*	18.20	7.2	15.00	5.9	2.4*
8–9	6 Pointer Adjustments	High	Low	41.9	5.5	39.5	5.2	2.2*	14.03	4.1	14.46	4.9	.5
10–11	3 Lights (Manual R.)	Low	High	46.2	5.8	46.3	4.8	.1	3.47	1.0	4.09	2.0	1.9
12–13	6 Pointer Adjustments	Low	High	41.7	6.5	42.5	4.6	.7	12.80	4.6	13.89	4.3	1.2

[a] Verbal reports after Trial 3 were scored on an arbitrary 6-point scale (high score = good performance). All other peripheral-task scores are cumulative reaction times in seconds (low score = good performance).

[b] The average correlation between the matching variable (score on the first two trials) and later performance on the central tracking task was .78. However, since the groups were not perfectly matched and since the matching variable was not effective in controlling individual differences on the peripheral task the t values given here are for unmatched groups. Corresponding t values for matched groups would be approximately twice as high as those shown for the central task.

* Significant at the 5% level.

** Significant at the 1% level.

were offered a bonus, the performance of Group I on the central task was significantly superior to that of Group II. Even on nonbonus Trials 4 and 7 Group I gave superior performance on the central tracking task. This is interpreted as indicating that the incentive continues to affect performance even on the interspersed practice trials, perhaps because *Ss* anticipate further bonuses. After Group II was offered a bonus for the first time (Trial 10), there were no further significant differences between the two groups on the central tracking task.

Within each group performance on the central task was significantly higher during bonus than during nonbonus trials.

The scores on the peripheral tasks were never improved by the bonus condition and were adversely affected in most instances. This result is in accordance with the initial hypothesis.

The first variation of the peripheral task was scored from 0 to 6 on the basis of the answers to the questions which followed Trial 3. One point was given for noticing any lights. Points were given for each of the three lights correctly identified, and an additional two points was credited to *Ss* who gave the correct sequence in which the lights were turned on. The difference in scores between groups is significant at the 5% level of confidence. Only 8% of *Ss* in Group II, the low-motivation group, failed completely to notice the peripheral lights on Trial 3, while 34% of *Ss* in Group I noticed no lights. This difference is highly significant.

During the second variation of the peripheral task (Trials 5–7) Group II had a shorter mean reaction time than Group I on all three trials, but the differences are not statistically significant.

The results of the third variation of the peripheral task (correcting for intermittent pointer deviations) were somewhat unexpected. During the initial practice trial on this task (Trial 7) Group I did significantly poorer than Group II, but during the next two (bonus) trials (Trials 8–9) the difference between the two groups was slight. The count of eye-fixations taken by the hidden *O* during these trials indicated that both groups had a mean of 24 eye movements between the central and peripheral task during 1-min. practice trials. During the bonus trials both groups showed fewer movements (longer fixations) but the number of changes in fixation for the high-incentive condition was greater (but not significantly so) than for the low-incentive condition. In spite of this greater frequency in checking the peripheral stimulus object, however, the group under high incentive did no better on this peripheral task than did the low-incentive group.

When incentives were reversed (Trials 10–13), the measures of performance on the peripheral tasks followed the predicted pattern. Group II now gave the poorer performance on the task of turning out the peripheral lights (Trials 10–11). The difference was significant at the 7% level. On the pointer-centering task the results also favored Group II, although the differences were not significant. On these trials, the relative number of changes in eye fixation for the two groups also reversed, being slightly larger for Group II than for Group I. The joint probability for the differences in the two peripheral tasks was significant at the 1% level, showing significantly poorer performance on peripheral tasks with high incentives following the reversal of incentive conditions for the two groups.

It seems clear that performance of the central tracking task was favor-

ably influenced by the high-incentive condition, while performance of the peripheral task was influenced adversely by this condition. The detrimental effects of the bonus upon reactions to peripheral stimuli were most pronounced when *Ss* had not been forewarned about the peripheral lights, less pronounced in the relatively simple light-reaction task, and least pronounced in the more exacting pointer-centering task.

Most of these results support the hypothesis that the efficiency of responses to peripheral stimuli is influenced by the degree of motivation, as well as by the difference in the expectation of reward from the central, as contrasted with the peripheral aspects of the task. While it may be assumed that the effect of the bonus upon motivation was approximately equal for all the tasks involved, the difference in expectation of reward was different for these three tasks. The lowest expectation of reward was present for the first variation of the perceptual task, noticing the lights, since *Ss* were not informed that the lights would go on or that lights had any significance. The expectation of reward from responses to the other two kinds of peripheral stimulation was greater, since *Ss* were instructed that their bonus would depend upon combined performance on both the central and the peripheral aspects of the task.

Several factors probably are involved in the smaller differences in performance on the third variation of the peripheral task. The *Ss* who were offered a high incentive looked at the instrument pointer more frequently during the 1-min. trial than did the low-incentive *Ss*. The third variation of the peripheral task actually had a greater effect on the combined task score than did the second variation, since it called for six (rather than three) responses, and required fine adjustive movements rather than gross reactions. Adjustment of the instrument pointer was actually a rather important aspect of the total task, and therefore, according to the hypothesis, should not have suffered greatly from high-incentive conditions.

Summary

The experiment tested the hypothesis that an increase in incentive results in increased perceptual selectiveness favoring those parts of the stimulus field which are interpreted by the *S* as most relevant to the expected reward.

A modified Buxton pursuit apparatus was used as a continuous central-tracking task. Three kinds of intermittent peripheral stimuli, differing both in the amount of relevance to reward expectations and in the facility with which they could be detected in peripheral vision, were employed.

Two incentive conditions were used. The low-incentive condition was produced by telling *Ss* that the trials were practice trials. The high-incentive condition was produced by offering a sliding-scale bonus, which ranged from 5 cents to 3 dollars, for good performance.

Results are in good agreement with the prediction that a condition of high incentive facilitates performance of a central task, but, in general, interferes with the performance of peripheral tasks. The detrimental results of the bonus upon responses to peripheral stimuli were greatest on the peripheral task in which expectation of reward can be assumed to have been the least.

(Received May 5, 1952)

References

1. Bahrick, H. P., Rankin, R. E., & Fitts, P. M. The effect of motivation upon peripheral perception during the performance of a central psychomotor task. San Antonio, Texas: USAF Human Resources Research Center, *Research Bulletin* 52–14, 1952.
2. Berlyne, D. E. Attention, perception and behavior theory. *Psychol. Rev.*, 1951, **58**, 137–146.
3. Bruner, J. S., & Goodman, C. C. Value and need as organizing factors in perception. *J. Psychol.*, 1947, **42**, 33–40.
4. Buxton, C. E., & Spence, K. W. Appraisal of certain tests of pilot aptitude. Washington, D. C.: Civil Aeronautics Administration Report No. 64, 1946.
5. Grether, W. F. A dual compensatory pursuit apparatus for use in psychological research. Dayton, Ohio; USAF Air Materiel Command Technical Report No. 60361, 1950.
6. Levine, R., Chein, I., & Murphy, G. The relation of the intensity of a need to the amount of perceptual distortion: a preliminary report. *J. Psychol.*, 1942, **13**, 283–293.

PLACING PRECISION AND ANGLE OF REGARD

ARTHUR J. RIOPELLE AND THOMAS M. STRITCH

Emory University

Any attempt to specify conditions for optimal performance on perceptual-motor tasks must consider the visual conditions under which performance is tested. The angle with which *S* views the work surface and whether he views it monocularly or binocularly may well affect proficiency. This experiment studies the effects of these variables as reflected in performance on a perceptual-motor task.

The pegboard task, requiring accurate and precise manipulations of small objects, is well suited for studying the effects of such conditions upon level of performance. Besides being simple to administer, the pegboard task also has the advantage of permitting *S* to reach a limit of ability within relatively few trials, thus emphasizing terminal performance rather than acquisition of skill.

METHOD

Apparatus.—The essential component of the apparatus was a pegboard mounted on an adjustable table. The pegboard was 14 in. long and 14 in. wide and contained four rows of ten pegholes each. The pegholes, spaced 1 in. apart, were approximately 3.5 mm. in diameter and would easily accept the pegs which were 3 mm. in diameter and 1 in. long. Spare pegs were stored in a 3-in. by 2-in. bin located 3.5 in. below the bottom row of pegholes.

Experimental design.—Eighteen viewing conditions were studied. Nine of the viewing conditions involved monocular vision whereas the remaining nine conditions involved binocular vision. The nine binocular conditions differed as to the angle with which the *S*s viewed the pegboard surface. The angles selected for study were 0, 5, 10, 30, 45, 60, 80, 85, and 90 degrees, all measured from the plane of the pegboard. An angle of zero degrees indicates that the plane of the pegboard was at eye level parallel to the line of sight. Identical angles were used when testing under monocular vision.

To each of the 18 viewing conditions was assigned a separate group of six *S*s. The members of each group were tested under only one viewing condition. They were drawn from the undergraduate student body at Emory University.

The experiment is thus arranged to form a two-row by nine-column factorial design. In addition to evaluating the significance of the two main effects, the significance of the interaction between these variables can also be determined.

Procedure.—Upon entering the laboratory *S* was seated before a table bearing the pegboard and was asked to sit erect but not rigid. He was further enjoined to keep the back of his head in contact with a headrest. The height of the eye was then determined and the table bearing the pegboard was lowered to the distance required to produce the desired angle of regard and adjusted so that *S*'s eyes were 400 mm. from the center of the pegboard. Goggles were then fitted to those *S*s viewing the pegboard monocularly. These goggles precluded vision with the left eye, but did not affect the vision of the other eye.

Standard instructions were read to *S* asking him to take one peg from the bin and to place it into the upper-left hole and to continue in this manner proceeding from left to right across the top row, then the second row, and so on. Twelve trials of 30-sec. duration were given separated by 30-sec. rest intervals. The *S* rested his hand on the near-right edge of the board during the rest interval.

RESULTS

The score on each trial for all *S*s was the number of placements. The mean number of placements per trial was computed for each group. These values are portrayed graphically in Fig. 1. This figure shows two curves, one for each viewing condition. The abscissa of the figure denotes the angle included between the plane of the pegboard and the elevation of *S*'s eyes. The ordinate represents mean number of placements per trial.

For both viewing conditions the level of performance increased with increasing angle of regard, reaching a

maximum when *S*'s eyes were directly above the center of the pegboard. In addition, the binocular performance was consistently superior to the monocular performance, although the superiority of performance is rather small over a major portion of the range of visual angles. Of special interest is the fact that monocular performance is markedly inferior at low angles of regard.

The interpretations outlined above were supported by the statistical analysis which revealed significant sources of variation due to angles of regard (.05 level) and viewing conditions (.05 level). The mean squares for these components were tested against the mean square for interaction which was itself significant at the .01 level of confidence when tested against the mean square for replication.

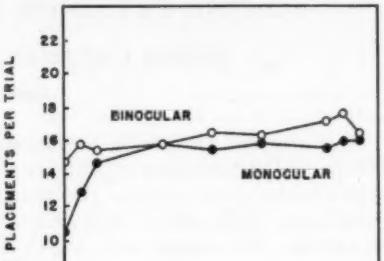


FIG. 1. Monocular and binocular performance on a pegboard test at varying angles of regard. The zero degree angle indicates that the eye was level with the plane of the pegboard, whereas the 90 degree angle indicates that the eye was directly above the pegboard.

DISCUSSION

From the results presented three facts can be noted. First, pegboard performance was significantly related to angle of regard, efficiency of performance increasing with larger angles of regard. Second, binocular performance was superior to monocular performance for all angles of regard, and third, this superiority was greatest for small angles of regard.

It is of interest to compare these results with the changes in optical stimulation which occur with changes in the angle with which a surface is viewed. Chief among these is the foreshortening of the surface with increasing slant. Just as the shadow of a circular object changes from a circular to an elongated, ellipsoid form on an increasingly slanting surface when illuminated by a single light source, so does the projection of a monocular "circle of confusion" change in form. This circle of confusion is measured in terms of visual angle from the axis of sight. It might be anticipated that the greater the projection of the circle of confusion, the poorer would be the localizability of a hole on the surface.

These same relations may be extended to the case of binocular vision. Under this condition fixation of a point results in intersecting projections of the circles of confusion. The important area of these projections is the overlapping region, and performance should be inversely related to this area. The area of this projection will also increase with increasing slant of the pegboard surface but it will not increase as rapidly as a single projection under monocular vision. Because of this relation binocular performance should be generally as good as monocular performance and decidedly superior to it at low angles of regard.

The circle of confusion, the limits of which are probably set by visual acuity, has been assumed to be con-

stant throughout all conditions of this experiment. However, at high angles of regard, the general superiority of the binocular groups may reflect the superiority of binocular over monocular acuity (2). The projection upon a textured surface of the circle of confusion, derived from visual-acuity considerations, however, may provide opportunities for other sources of information about localization to arise and such cues may restrict the effective region of uncertainty. Gradients of stimulation from the texture of the surface (1), horizontal and vertical head and eye movements, memory and kinesthetic cues may well be important factors in determining proficiency of localizing responses. That these factors may assume predominant roles under appropriate circumstances is suggested by the ability of our *Ss* to make *some* responses at zero angle of regard even though optical stimulation from the pegboard surface theoretically ceases.

Summary

Eighteen groups of *Ss* were tested on a laboratory-constructed pegboard test. Nine of the groups viewed the pegboard monocularly and the remaining nine groups viewed the pegboard binocularly. The nine groups of each viewing condition were tested at nine angles of regard varying from zero to 90 degrees, one group at each angle.

For angles of regard greater than about 5 or 10 degrees, slight but consistently increasing proficiency of performance was found. At all angles of regard the binocular performance was superior to the monocular performance. Furthermore, the difference in performance between the two viewing conditions increased at very small angles of regard.

The results are interpreted in relation to the changes in optical stimulation accompanying variations in the angle of regard.

(Received May 5, 1952)

References

1. Gibson, J. J. *The perception of the visual world.* New York: Houghton Mifflin, 1950.
2. Horowitz, M. W. An analysis of the superiority of binocular over monocular visual acuity. *J. exp. Psychol.*, 1949, **39**, 581–596.

EFFECT OF DISCRIMINATION REVERSAL ON HUMAN DISCRIMINATION LEARNING

RICHARD D. WALK[1]

Arlington, Virginia

These experiments are concerned with a test of the continuity hypothesis, using human beings as *S*s. Animal experimentation has given support for the gradualness of the learning process, as continuity theory (**15**) maintains. Such research has shown that discrimination reversal, while the animals are responding in a chance fashion to the later-to-be-learned cues, definitely slows up learning (**3, 5, 13, 18**). The fact that discrimination reversal slows up learning does not support the discontinuity view (**9**) that the animal learns by testing hypotheses and, while testing an incorrect hypothesis, learns nothing of the finally-to-be-learned discrimination.

Only one experiment had been published to test the continuity hypothesis on human *S*s. Prentice (**12**) reversed one group after 20 trials and compared it with a group that had not been reversed. He secured continuity results. The reversed group was approximately 20 trials slower, in reaching the learning criterion, than was the nonreversed group. The 20-trial difference, however, is the amount of reinforcement the reversed group received on the opposite principle to the one finally learned. Prentice interpreted his results as favoring neither a continuity nor discontinuity position since there was no decrement beyond the 20-trials reinforcement the reversed group received on the opposite principle. This is a misinterpretation of the original continuity-discontinuity controversy which followed from Lashley's assertion that ". . . the actual association is formed very quickly and . . . both the practice preceding and the errors following are irrelevant to the actual formation of the association" (**10**, p. 135). In other words, the correct hypothesis must appear suddenly; and the principle reinforced in preliminary training is of no consequence as long as the organism is not testing the correct hypothesis at the time of reversal. Prentice used verbal report as a measure of "insight" and discarded from both groups those *S*s testing the correct hypothesis ("insight") before Trial 20. According to discontinuity theory, there should have been no difference in the speed of learning of his two groups.

Prentice only presented the number of trials for each group to reach solution. He presented no group-learning curves. Our experiments will present group-learning curves in order to determine whether there is a statistically significant decrement following reversal for a reversal group as compared to a nonreversal group. Such a decrement will be proof that learning has been taking place during the period prior to reversal and, hence, support the continuity view of discrimination learning.

In addition to determining whether a decrement follows reversal for the experimental group we shall be inter-

[1] This report is a portion of a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree in the Department of Social Relations, Harvard University. The author wishes to acknowledge the direction of Dr. Richard L. Solomon in the conduct of this research. This research was supported by the Laboratory of Social Relations, Harvard University.

ested in the following: (*a*) the effect of reversal on number of trials to reach correct solution; (*b*) the number of trials the decrement hypothesized for the experimental group lasts as compared to the control group; (*c*) the effect of discrimination reversal on group latencies; (*d*) hypothesis (verbal) behavior in the experiment.

Procedure

Apparatus.—The *S* faced a black cardboard panel, 28 in. high and 22 in. wide, which rested on an ordinary table. In the center of the panel, 13½ in. up from the bottom, was cut an aperture 7 in. long and 3 in. wide. A screen controlled by *E* could be lifted to expose two stimulus cards side by side in the aperture. Beside the aperture were two lights, one colored green ("right") and one red ("wrong"). Buttons encased in hollow wooden handles were on the table and beside the hands of *S*. The button beside the right hand of *S* corresponded to the right-hand side of the apparatus, and vice versa. If the correct stimulus were in the right-hand side of the apparatus, for example, and *S* pushed the left button, the red light "wrong" would be illuminated. Choice of a stimulus by pushing one of the buttons also stopped an electric clock that was started when the aperture was opened. The *E* lowered the screen 1–2 sec. after *S* made his choice. The apparatus, thus, served two functions. First, both *E* and *S* learned immediately whether each choice was correct or incorrect. Second, the latency of *S*'s response was available to *E*. While differing in details, the apparatus is grossly similar to those used by Buss (2), Marquardt (11), and Prentice (12).

Stimulus presentation.—The order in which each side of the aperture was correct was determined by an order of restricted randomization. It was taken from a table of random numbers with the limitations that (*a*) one side of the aperture could be correct no more than four times in succession, and (*b*) each side of the aperture was to be correct ten times out of every 20 trials. Three different orders of stimulus presentation, each of 100 trials in length, were prepared. The same order of stimulus presentation was used in both experiments.

In neither experiment were stimulus cards paired together at random. Rather, stimuli were paired together only when they differed in at least two characteristics in addition to the discriminating one. This was to prevent *S* from attaining a quick solution because, if presented randomly, stimuli that were almost identical except for the discriminating characteristic might have appeared together.

Instructions.—No prearranged instructions were read to *S*. The *S* sat down facing the panel and *E* opened the aperture exposing two sample stimuli different from those used in the experiment. In the course of his exposition, *E* tried to make clear that the "concept" *S* was seeking could be simple or complex, that it could be a relationship involving either the positive or the negative stimulus, and that it could depend on relative characteristics, absolute characteristics, the side of the aperture, or a combination of all of these. The *S* was also asked to verbalize his hypotheses from time to time because *E* was interested in "the process by which people solve problems." The *S* was told to take as long a time on each trial as he wished, and also that the experiment would end if he verbalized the correct concept and, in addition, attained a certain number of correct responses.

Experiment I

Purpose.—This experiment was designed to test some specific implications of the continuity hypothesis in a discrimination problem which, it was known from preliminary experimentation, *no* S *would "solve."* However, it had been observed that *S*s seemed to get more correct responses than chance, partly as a result of memory, and partly as a result of unverbalized "knowledge" of the underlying discriminable difference. It was therefore hypothesized that preliminary training on the same principle as that used in the final learning task should lead to some positive transfer effect. Furthermore, preliminary training on a principle opposite to that used in the final learning task should lead to a decremental effect on final learning. A control group, with no preliminary training, should serve as a check on how much is learned in preliminary training. Under these conditions, any decrement in the group with the reversed training, as compared to the other groups, would support the continuity hypothesis.

Stimuli.—The stimuli were painted wooden blocks that varied in size, shape, and color. They were cemented in the center of 3 × 3½-in. white cardboard stimulus cards. The stimulus component chosen for reinforcement was an absolute height of ¾ in., while heights of ⅝ and 1⅛ in. were "wrong." A discriminable distance of ½ in. separated one height from the other. With this exception, the patterns in the blocks varied approximately at random. There were shapes, colors, and depths that were similar at each height, though notches and bevels were cut in blocks to make every shape discriminably different. Hereafter, stimuli with a height of ¾ in. will be referred to as *A-blocks*, and those with a

TABLE 1

STIMULUS CONDITIONS IN EXPERIMENT I

Group	Condition	Preliminary Training (100 Trials)	Final Learning Task (200 Trials)
I	Positive transfer	A-blocks positive (100 different pairs)	A-blocks positive (50 pairs)
II	Negative transfer	B-blocks positive (100 pairs)	A-blocks positive (50 pairs)
III	Control	None	A-blocks positive (50 pairs)

height of $\frac{3}{8}$ or $1\frac{3}{4}$ in. will be referred to as *B-blocks*.

One hundred and fifty different pairs of stimuli were used. One hundred pairs of stimuli were used for preliminary training for Groups I and II. The other 50 pairs of stimuli were used by all groups for the final learning task.

Subjects.—The *S*s were 30 students from Harvard University and Radcliffe College.

Experimental design.—The experimental design is summarized in Table 1. One hundred trials were given to *S* in a 1-hr. session, and no more than one experimental session was held per day. The *S*s in Groups I and II had three 1-hr. experimental sessions and *S*s in Group III had two 1-hr. sessions.

Since each *S* saw the 50 pairs of stimuli used in the final learning task four times, there was the possibility of memory for specific stimuli. The possibility of serial memory effects was controlled by starting each block of 50 trials at a different point in the stimulus order (with stimulus pairs No. 1, 13, 26, 37). A-blocks were also paired with two different B-blocks. The memory problem, however, is secondary since differences were predicated on the basis of the preliminary training.

Experiment II

Purpose.—This experiment serves as a check on the results of the first experiment. This experiment used an *easily solvable problem*, and hence we could test not only for a decrement in the reversal group as compared with a nonreversal group but also test for the effect of reversal on the number of trials required to reach a correct solution. "Social" stimuli, photographs of a girl student, were used; and thus the extent to which "expressive behavior" components entered into *S*'s orientation to the discrimination-learning task could also be investigated.

Stimuli.—Thirty-six different photographs of a Radcliffe undergraduate were used to construct the stimuli for this experiment.[2] She posed for 24 different face and head positions by systematically varying head, eye, and mouth poses. These poses were: *head* (left tilt, straight, right tilt), *eyes* (open, closed) and *mouth* (open, closed, smiling, not smiling). This is, thus, 3 (head) × 2 (eyes) × 4 (mouth) or 24 head positions. The number of trunk positions was 10, by varying palm and hand positions. These were: *palms* (facing camera, facing rear) and *hands* (by side, horizontally out to side, halfway over head, facing front at waist level, facing front at shoulder level). These make 2 (palm) × 5 (hand) or 10 trunk positions. The *E* sliced the photographs approximately at shoulder level and combined the 24 head positions with the 10 trunk positions, making 240 different stimulus photographs. Sample stimuli are illustrated in Fig. 1. The photographs measured 4 × $3\frac{1}{4}$ in. It was necessary to enlarge the aperture in the panel of the apparatus from a height of 3 in. to one of 4 in. for this experiment.

The concept of "mouth open, not smiling" was chosen to be of medium difficulty, and it appeared in 60 stimulus cards. Sixty negative stimuli were selected from the other mouth positions: "smiling, mouth open" (20 stimuli), "smiling, mouth closed" (17 stimuli), "not smiling, mouth closed" (23 stimuli). The other 120 photographs were not used in the experiment.

Subjects.—The *S*s were 60 students from Harvard University and Radcliffe College.

Experimental design.—For a balanced design, "mouth open, not smiling" was used both as a correct and as an incorrect concept. That is, half the *S*s were reinforced for choosing "other mouth positions" as the correct stimuli. All *S*s who had insight on Trial 20 or before were discarded for the test of the continuity hypothesis. A summary of the experimental design is presented in Table 2.

No *S* was retained in the experiment for a longer period than two 1-hr. sessions of 60 trials each (120 trials). Since discrimination reversal would be used, for some *S*s, after 20 trials, stim-

[2] The author wishes to thank Miss Elizabeth Bibber, Class of 1952 of Radcliffe College, for posing for the photographs from which stimuli were made.

FIG. 1. A sample pair of stimulus photographs, showing "mouth open, no smile" compared to "mouth open, smiling"

ulus pairs 1–20 were only shown to *S* once. The stimulus presentation order was:

Trials 1–60	Stimulus Pairs 1–60
Trials 61–100	Stimulus Pairs 21–60
Trials 101–120	Stimulus Pairs 21–40

Of course, once *S* verbalized the correct hypothesis and achieved ten correct choices in succession, the experiment was stopped. In this experiment, those *S*s who got 23 out of 25 correct choices and had an hypothesis that was nearly correct were also stopped.

RESULTS

Experiment I.—An inspection of Fig. 2, where learning curves are plotted, shows that Group I is slightly superior to Groups II and III over the initial 25 trials of the final learning task. Group II, the reversal group, is apparently inferior to both Groups I and III for the initial 25 trials, but this decrement does not last any longer. The initial data, then, are in accordance with a continuity prediction; but an inspection of Fig. 2 also shows that the transfer lasts no longer. In Table 3 the means and *SD*'s for number of correct responses for all groups on Trials 1–50, and on Trials

TABLE 2

STIMULUS CONDITIONS FOR EXPERIMENT II

$N = 15$ for each group

Group	Positive Stimulus	Negative Stimulus
I-A Straight learning	Mouth open, no smile	Other mouth positions
I-B Straight learning	Other mouth positions	Mouth open, no smile
II-A Discrimination reversal		
1. Trials 1–20	Other mouth positions	Mouth open, no smile
2. Trials 21–*n*	Mouth open, no smile	Other mouth positions
II-B Discrimination reversal		
1. Trials 1–20	Mouth open, no smile	Other mouth positions
2. Trials 21–*n*	Other mouth positions	Mouth open, no smile

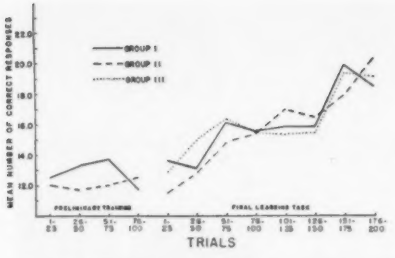


FIG. 2. The mean number of correct responses during preliminary training and the final learning task: Exp. I

51–200 of the final learning task, are presented. For Trials 1–50, Group II is inferior to Group I (with a non-significant t of 1.1, however), and the control group is superior to both. Transfer, if there is any, lasts no longer than the first 50 trials. After Trial 50 all groups are similar in performance, presumably because they have seen all of the stimuli once and should now be able to remember some of them. From Trial 51–200 all groups are similar in the number of correct responses attained.

Cumulative correct responses, by one-trial intervals, were also plotted for Groups I and II, for Trials 1–50. The decrement for Group II started at about Trial 5 and was maximally operative around Trial 20. At this time the t value (Trials 1–20 for each group compared) is significant beyond the .05 level. Our prediction, however, is only anchored at Trial 1 by our experimental operations. We had not predicted Trial 20 as the time of maximal decrement.

TABLE 3

MEAN CORRECT RESPONSES ON THE FINAL LEARNING TASK: EXP. I

Group	N	Trials 1–50		Trials 51–200	
		Mean	SD	Mean	SD
I	10	26.6	3.6	101.4	14.4
II	10	24.3	5.5	101.4	18.5
III	10	27.7	2.3	100.7	10.8

Mean log latencies for Groups I and II, whose experience in the experiment was similar, were computed for Trials 1–30 of the final learning task. Latencies of the two groups did not differ significantly. Reversal, apparently, does not change the response latencies.

No *S* gave any indication in his verbalized hypotheses that he was aware of reversal. Some *S*s reported being able to pick the correct stimulus figure without being able to define the reason for it. This "learning without awareness" was reported in preliminary experimentation and was the reason we continued experimentation even though *S*s could not learn the concept. Ten of our 30 *S*s reported using "intuition," "vague feelings," etc. that may let us infer some degree of "learning without awareness." A few sample hypotheses are given:

Ba: "I just had a vague feeling for them."
Wo: "Just sat back and let an impression trickle through. This was the most successful."
Ho: "It is the strangest thing the way these develop an atmosphere of 'rightness' and 'wrongness' without any specific reason."

Experiment II.—In this experiment no *S* who had "insight," that is, who was testing an hypothesis concerning mouth or facial expressions, was retained in the experimental test of the continuity hypothesis. Despite insight, however, experimentation continued, though no "insightful" *S*s were reversed. Table 4 shows the number in each group who had "insight" and the number who did not. Approximately two-thirds of the *S*s in each group did not have insight and could be used to test the continuity hypothesis. The reader will note that

TABLE 4

THE NUMBER OF *Ss* WHO HAD "INSIGHT" DURING TRIALS 1–20, AND NUMBER OF *Ss* REMAINING WHO SOLVED THE PROBLEM

Group	*N*	*Ss* Showing "Insight" during Trials 1–20	*Ss* Remaining after Trial 20	*Ss* Solving Problem after Trial 20
I-A	15	6	9	3
I-B	15	5	10	4
II-A	15	4	11	9
II-B	15	6	9	2
Total	60	21	39	18

not all "insightful" *Ss* attained enough correct responses to stop the experiment.

Following reversal, Group II showed a definite decrement as compared with Group I. In Table 5 is presented the number of individuals in each group who are above and below chance for Trials 21–30. Fisher's (4) exact test, to determine whether the reversal group is significantly different from the straight-learning group, yields a one-tailed probability of less than .005. Learning curves are presented in Fig. 3 for all groups. It can be seen that both reversal groups (II–A and II–B) are inferior to the straight-learning groups (I–A and I–B) for Trials 21–30. It should be remembered that the only *Ss* considered are the ones who were *not* working on a component of the correct solution when they had reached Trial 20.

TABLE 5

NUMBER OF CORRECT RESPONSES ON TRIALS 21–30 FOR *Ss* WHO DID AND DID NOT HAVE DISCRIMINATION REVERSAL ON TRIAL 21

Group	Number of Correct Responses*		
	4 or Less	5 or More	Total
Straight learning	6	13	19
Discrimination reversal	16	4	20
Total	22	17	39

* No *Ss* included who had "insight" during Trials 1–20.

As in the previous experiment, cumulative correct responses were plotted trial by trial for both groups. The decrement of Group II as compared with Group I lasted approximately 10–12 trials. Mean log latencies for both groups were also plotted, and latencies did not change as a result of the reversal experience.

No data can be presented on the effect of reversal on the attainment of a correct solution. This is because

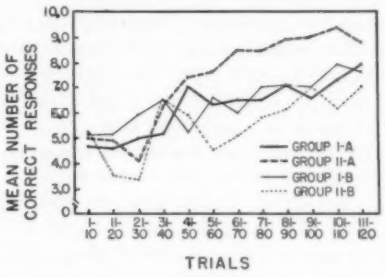


FIG. 3. The mean number of correct responses during training: Exp. II

our two subgroups (I–A and I–B; II–A and II–B) were not equivalent. A reinspection of Table 4 shows that 9 out of 11 *Ss* in Group II–A, a reversal group, solved the problem. This is a larger proportion than any other group. Did reversal introduce a "disinhibiting" effect which focused attention on the correct stimulus component? Blum and Blum (1) mention this as a possibility. The reason why we cannot infer this, and the reason why we cannot compute the effect of reversal on the attainment of a solution, is an interesting one.

In Table 2 we gave the stimulus procedures for each subgroup. The

TABLE 6

NUMBER OF *Ss* WHO ATTAINED THE SOLUTION

Stimulus Condition	Solution	No Solution	Total
Mouth open, no smile	23	9	32
Other mouth positions	10	18	28
Total	33	27	60

subgroups "A" had "mouth open, not smiling" as a positive stimulus (after Trial 20) while those subgroups "B" had "mouth open, not smiling" as the negative stimulus. The positive stimulus for the "B" groups was *three* mouth positions. All groups had the same stimuli, as we mentioned before; only the reinforced stimulus differed.

The results seem to indicate that having "mouth open, not smiling" as a positive stimulus is not the same as having it as a negative stimulus. In Table 6 we have presented the data, for all our *Ss*, on those who solved or did not solve the problem, comparing those with "mouth open, not smiling" as a positive stimulus with those with "other mouth positions" as a positive stimulus. Many more solutions were attained by *Ss* with "mouth open, not smiling" as the positive stimulus. To put the data in the form of Table 6, we considered the first 20 trials as separate from the trials beyond Trial 20. An *S* with "insight" by Trial 20, it will be remembered, was treated as if he were in a straight-learning group. Without insight, an *S* either stayed in the same group (straight learning) or went into a new group (discrimination reversal). Since two more *Ss* in the "mouth open, not smiling" groups had "insight," the final number of *Ss* is 32 for this group as compared with 28 for the other group. By using Fisher's exact test on the data of Table 6 the probability is .0103 (both-tails).

One might ask, since 9 out of 11 in reversal Group II–A solved the problem, whether these results might be an artifact of the disinhibiting effects of reversal. We do not believe this to be the case because Group II–B, which also had discrimination reversal, had only two out of nine solve the problem (see Table 4).

Further support for the difference of the "mouth open, not smiling" as a positive as compared to a negative stimulus comes from the number of *Ss* in each group with "insight" who achieved enough correct responses to stop the experiment. There were 12 *Ss* who had "insight" while "mouth open, not smiling" was a positive stimulus. Eleven of these solved the problem. Of the nine *Ss* with "insight" with "other mouth positions" as a positive stimulus only four attained enough correct responses to stop the experiment.

Verbal hypothesis behavior also supports our contention that "mouth open, not smiling" differed as a positive and as a negative stimulus. "Other mouth positions" had more partial solutions as illustrated by the following protocols:

Zo: "When she is smiling it is right. Otherwise, I take arms bent so closer to face, but this does not work all the time."

Cu: "Smiling one always picked. If not smiling it depended on foot position. . . ."

Ga: "Mouth had to be shut. When both mouths open, you just guessed."

From this analysis, then, it appears that we have four experimental groups instead of two. This precludes any intensive comparisons. To factor out the influence of reversal on the attainment of solution as well as the influence of "positive" and "negative" stimuli would require a new experiment.

Only one *S* gave a hypothesis that seemed to show some "awareness" of the reversal relationship. He was reversed from "other mouth positions" to "mouth open, no smile." He showed practically no decrement, being in the group with five or more correct responses in Table 5, and he soon got all choices correct. After stating the correct hypothesis, he said, "you know for a while I steered away from the open mouth." Even this was not completely correct because mouth "open, smiling" was correct before reversal. The only *S* who gave a hypothesis showing he may have had some awareness of discrimination reversal is in the discrimination reversal, 5 or more, category of Table 5.

As in our previous experiment, *Ss* verbalized

about using "intuition" or "feelings" to aid in discriminative behavior. One *S*, well above chance, remarked, "Whenever I analyze it, I lose it." (He later expressed the correct hypothesis.) Another *S* reported, "That bunch I got right without realizing why. When I got one wrong, it disturbed me and I realized it was because both mouths were open."

We would expect that many *S*s would use the "expressive behavior" of the photographs to orient their choices. A few sample protocols follow:

Pie: Expression of face. Either she has a happy expression with a smile or a contented expression with eyes closed. When mouth open, it is not right as long as no contentment expressed.

En: The other face looks happier. This one has a baffled, wornout look.

Cr: Trial 10: I lean toward the one that is least vivacious, the one in a trance. (Insight, Group I–A.)
Trial 90: Narrowed it down to concentrating entirely on facial expression. Most disjointed expression—I can't seem to put my finger on a specific trait. (14 choices correct at this point.)
Trial 105: In borderline cases, I stick to imbecility rather than repose. (29 correct in succession.)

RG: Right one seems to be presenting a melancholy song or poem and wrong one is on the flirtatious or gay side.

Discussion

We shall discuss the following results in the two experiments: (*a*) the decrement following discrimination reversal, (*b*) "learning without awareness," and (*c*) the nonequivalence of the four experimental groups in Exp. II.

Our experiments seem to demonstrate that, with complex stimuli, cue reversal has a decremental effect on discrimination learning, therefore supporting the continuity hypothesis. But this effect is not of long duration.

How can we account for these results? The initial decrement favors the continuity position, but the absence of long-range effects does not support the continuity side. One difference between our human experimentation and previous animal experiments concerns the number of stimulus components. When the discrimination is reversed for the rat, for example, the only stimulus component reversed is usually a fairly simple one. Black becomes white (**18**), for example, or horizontal patterns become vertical patterns (**13**). Our experiments have reversed far more than one stimulus component. In Exp. I, a height relationship was completely reversed but certain characteristics that concern color, shape, depth, and size were a large part of the total stimulus pattern. Random variations were continually "reversing" some aspects of the other stimulus components. Likewise, in Exp. II mouth positions are only a few of the varying stimulus components. Arms, hands, eyes, shadows, and other unknown components are always shifting in emphasis. Reversal of the dominant or correct stimulus component is only an accentuation of what is happening continually to other stimulus components. The effect, although definite, soon becomes mixed in with the welter of other stimulus dimensions.

Another possible explanation comes in terms of the work of Spence (**15**) and Harlow (**6**). If Spence's position is correct, and the tendency to respond to a stimulus is dependent on the acquisition of excitatory strength to the positive stimulus, we may posit that, in the experienced organism, excitatory tendencies to all stimulus components are initially at a fairly high level. The learning task cannot build up a large initial increment. Reversal, then, does not make for a large proportionate change in excitatory strengths. The work of Harlow supports this type of interpretation. Reversal did not have large effects on his *S*s, though a complicating factor was that his *S*s were reversed *after* they had learned a task and had "learned" to reverse their habits. Spence's position, then, might predict that the larger the initial excitatory strength to all stimulus components, the smaller will be the total decrement during presolution reversal. Reversal would *always* have a decremental effect, however.

These two factors, the presence of a large number of stimulus components and the use of experienced organisms, may help explain the initial decrement together with a lack of long-range effects. These results are in accordance with continuity theory, but also suggest that continuity theory might be broadened theoretically to account for the type of situation we have used here.

A second finding in our experiments, the presence of "learning without awareness," is not novel. The experiments were designed to test the continuity hypothesis and not to examine learning without awareness. The fact that so many of our *S*s reported being able to choose the correct stimulus without being able to define the reason for it, led us to include some of their verbal hypotheses in our results. If *S*s are reversed

and show a decrement in learning in comparison with another learning group, yet show no awareness of reversal in their hypotheses, they must have learned without awareness. The decrement following reversal seems to be proof that (*a*) they have learned, and (*b*) the absence of a verbal report denoting the reversal shows that this learning was without awareness.

Our last result, the nonequivalence of the four experimental groups in Exp. II, deserves comment because of its implications for future research. The subgroups were not equivalent because, apparently, "mouth open, not smiling" was more effective as a *positive* stimulus than was a positive stimulus of three mouth positions when "mouth open, not smiling" was the *negative* stimulus. This experiment seems to show the importance of the rewarded stimulus as compared to the negative stimulus. Broadly speaking, this is in accord with stimulus-response reinforcement theory, with its emphasis upon reward; but attempts to analyze excitatory and inhibitory potential do not make a direct statement on the question. Hull (**8**) and Spence (**16, 17**) both make the inhibitory gradient in discrimination learning much weaker than the excitatory gradient. However, excitatory strength can be formed through learning contributed by either stimulus; that is, directly through reward or, indirectly, by learning to avoid the negative stimulus. Hilgard (**7**) made an algebraic analysis of conditioned eyelid discrimination in humans. He solved equations to determine the influence of reinforcement, nonreinforcement, and voluntary positive and negative tendencies. The data gave negative tendencies such a small contribution that Hilgard mentioned the possibility of a strong initial negative tendency to conditioning that had to be overcome.

The work of Smoke (**14**) on positive and negative instances in concept learning is a related line of investigation, but not directly comparable to ours. Smoke let *S* know in advance which stimuli contained the correct concept and which did not. For simultaneous presentation, for instance, a large number of examples of the concept might be spread before *S*. His task would be to study the stimuli until he thought he knew the concept. Under these conditions, *S*s found negative instances an aid in simultaneous presentation but confusing when presented successively. In our experiment the negative instance appeared to be less helpful than the positive instance.

Summary

Two experiments were designed to test the continuity hypothesis of discrimination learning in human *S*s. Experiment I consisted of a problem which *S*s could not solve. The *S*s with preliminary training to (*a*) the same principle, and (*b*) a principle opposite to that on a final learning task, were compared to a control group with no prior training. In Exp. II photographs of a girl were used as stimuli, and *S*s solved a problem concerning facial expressions. One group was reversed after 20 trials on the opposite principle. Only *S*s without "insight" at the end of 20 trials are considered in the test of the continuity hypothesis.

The results are as follows:

1. In both experiments a decrement for the reversal group was observed following discrimination reversal.

2. In neither experiment was the decrement following reversal of long duration.

3. In neither experiment were group latencies affected as a result of reversal, nor were *S*s aware of reversal, as evidenced in their verbalized hypotheses.

4. In Exp. II the positive stimulus appeared to be much more important than the negative stimulus. This is interpreted as being in agreement with general stimulus-response reinforcement principles.

5. The *S*s reported the phenomenon of "learning without awareness" in both experiments.

6. It is suggested that the decrement following reversal, with apparently no long-range effects, can be interpreted theoretically by taking account of (*a*) the number of stimulus components in a complex human visual discrimination, and (*b*) the background of experience of the human organism in discrimination learning.

(Received May 5, 1952)

References

1. Blum, R. A., & Blum, J. S. Factual issues in the "continuity" controversy. *Psychol. Rev.*, 1949, **56**, 33–50.
2. Buss, A. H. A study of concept formation as a function of reinforcement and stimulus generalization. *J. exp. Psychol.*, 1950, **40**, 494–503.
3. Ehrenfreund, D. An experimental test of the continuity theory of discrimination learning with pattern vision. *J. comp. physiol. Psychol.*, 1948, **41**, 408–422.
4. Fisher, R. A. *Statistical methods for research workers.* (10th Ed.) London: Oliver and Boyd, 1946.
5. Gatling, F. P. A study of the continuity of the learning process as measured by habit reversal in the rat. *J. comp. physiol. Psychol.*, 1951, **44**, 78–83.
6. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.
7. Hilgard, E. R. An algebraic analysis of condioned discrimination in man. *Psychol. Rev.*, 1938, **45**, 472–498.
8. Hull, C. L. The problem of primary stimulus generalization. *Psychol. Rev.*, 1947, **54**, 120–134.
9. Krechevsky, I. A study of the continuity of the problem-solving process. *Psychol. Rev.*, 1938, **45**, 107–133.
10. Lashley, K. S. *Brain mechanisms and intelligence.* Chicago: Univer. Chicago Press, 1929.
11. Marquardt, D. I. The pattern of punishment and its relation to abnormal fixation in adult human subjects. *J. gen. Psychol.*, 1948, **39**, 107–144.
12. Prentice, W. C. H. Continuity in human learning. *J. exp. Psychol.*, 1949, **39**, 187–194.
13. Ritchie, B. F., Ebeling, E., & Roth, W. Evidence for continuity in the discrimination of vertical and horizontal patterns. *J. comp. physiol. Psychol.*, 1950, **43**, 168–180.
14. Smoke, K. L. Negative instances in concept learning. *J. exp. Psychol.*, 1933, **16**, 583–588.
15. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.
16. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.
17. Spence, K. W. Failure of transposition in size discrimination of chimpanzee. *Amer. J. Psychol.*, 1941, **54**, 223–229.
18. Spence, K. W. An experimental test of the continuity and non-continuity theories of discrimination learning. *J. exp. Psychol.*, 1945, **35**, 253–266.

REMINISCENCE IN A MANIPULATIVE TASK AS A FUNCTION OF WORK-SURFACE HEIGHT, PREREST PRACTICE, AND INTERPOLATED REST

DOUGLAS S. ELLIS

Iowa State College

VICTOR MONTGOMERY AND BENTON J. UNDERWOOD

Northwestern University

Recent research on reminiscence in motor tasks indicates that amount of reminiscence is proportional to (*a*) length of interpolated rest (**2, 10, 15**), (*b*) number of prerest responses (**2, 10, 14**), and (*c*) degree of massing of prerest practice (**3, 13**). Result (*a*) suggests that reminiscence can be attributed to the dissipation over rest of a performance depressant: results (*b*) and (*c*) suggest that this depressant is generated during prerest practice. A hypothetical performance depressant (reactive inhibition, I_R) with these characteristics has been postulated by Hull (**8,** pp. 277–303) and summarized by a number of authors (**12, 15, 18, 20, 21**). I_R is defined as a response-produced negative motivational state which has the property of dissipating spontaneously with time. The amount of I_R generated by a sequence of responses is assumed to be directly proportional to both the number and effortfulness of responses. From such theory it would be expected that increases in task effortfulness should increase reminiscence.

The pursuit-rotor experiment of Eckstrand (**4**) tests this prediction. Eckstrand varied effortfulness by requiring *S* to depress the stylus against varying amounts of spring tension, and failed to find significant differences between effortful conditions in either prerest performance or reminiscence. He attributes the prerest results to greater habit strength offsetting greater inhibition at the more effortful conditions, and suggests that conditioned inhibition may have deflated reminiscence measures from these conditions.

The present experiment tests the prediction of proportionality between response effort and reminiscence for the case where effort is varied by manipulating work-surface height. In a previous experiment (**6**) it was found that a high work-surface height was most decremental to performance, while optimal performance occurred at a moderate height. If the high-height performance decrement is due to I_R, greater reminiscence should occur at the high height than at the moderate height. Amount of prerest practice and length of rest period are included to trace the generation and dissipation of I_R at the two heights.

METHOD

Experimental design.—The factorial design incorporated the following levels of the three independent variables: (*a*) *Work-surface height:* 3 in. below (moderate height) and 8 in. above (high height) *S*'s elbow height; (*b*) *Amount of prerest practice:* 2, 8, and 16 min. of massed practice; (*c*) *Rest-period length:* 2 min., 10 min., and 24 hr.

Subjects.—The use of 180 *S*s provided ten replications of the 18 experimental conditions required by such a design. The *S*s were male volunteers from the introductory psychology course at Northwestern University during the winter and spring of 1950 and were randomly assigned to height conditions. However, they were assigned to practice-rest conditions within

heights so that these groups were comparable in initial ability.

Apparatus and task.—Work-surface height was controlled by the adjustable work surface used in previous research (6). Mounted on the work surface were four eight-block rows of the Minnesota Rate of Manipulation Test. Gross postural shifts were minimized by requiring *S* to keep his feet within an 18 × 18-in. area delimited by 3-in. boards. Lighting was provided by a fluorescent lamp attached to the work surface.

The *S* was required to use both hands in turning the blocks over in the holes. The task was performed by rows, with the direction of motion on a row systematically varying from left to right as *S* worked from the outer to the inner rows. When the last block on the inner row was turned, *S* reversed directions on that row and worked from inner toward outer rows. The *S* was instructed to avoid fumbling the blocks over and was not to rest his elbow on the work surface while performing the task.

Procedure.—To avoid differential amount-of-work sets (16, 19), *S*s were not told the length of either practice or rest periods. Following brief preliminary instructions and a demonstration of desired work methods, *S* was given a 64-block practice trial at the appropriate height. During this trial improper work methods were corrected. The *S*s were then instructed that the experiment involved "quite a bit of work" and were given an opportunity to withdraw. No *S* took advantage of this opportunity.

During the ensuing massed practice, *S* was under instructions to work as fast as possible until *E* signalled a stop. Knowledge of results (blocks turned per minute) was given every minute. After completion of prerest practice, *S*s in the 2- and 10-min. rest conditions rested in an outer room equipped with current magazines. The *S*s receiving the 24-hr. rest were merely instructed to return to the laboratory at the appropriate time.

At the close of the rest, all *S*s were told that five more minutes of massed practice with knowledge of results would be required. Upon completion of this postrest practice, *S* was instructed not to discuss the experiment with other students.

Results [1]

Basic data.—Prerest practice curves at the two heights are presented in Fig. 1. The plotted points are performance means utilizing all available data: minutes 1–2 are based on all 180 *S*s (90 at each height); minutes 3–8 are based on the 120 *S*s (60 at each height) who received either 8 or 16 min. of prerest practice; minutes 9–16 are based on the 60 *S*s (30 at each height) who received 16 min. of prerest practice. The linear equations are least-squares fits to these data.

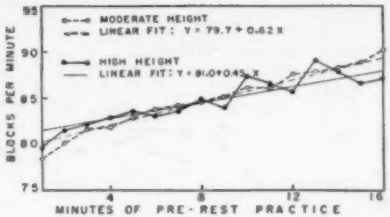


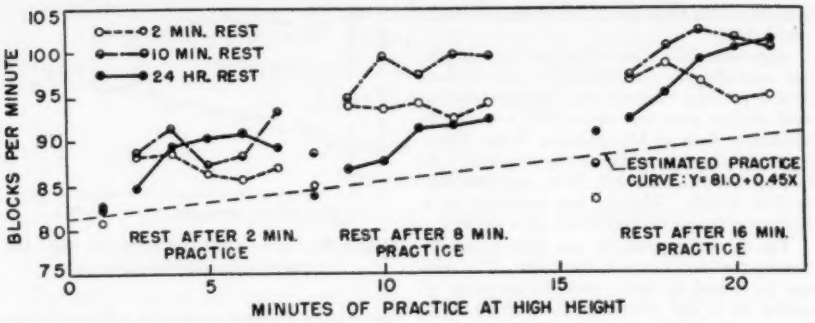
FIG. 1. Prerest practice curves at the high and moderate work-surface heights

Although the high-height curve is considerably more variable in the later stages of practice,[2] prerest performance level is comparable for the two heights. Analyses of variance performed at 2, 8, and 16 min. of practice indicate that no significant differences exist between performance at the two heights. These analyses also show that the various practice-rest groups within height conditions are comparable in prerest performance.

Postrest curves are plotted in Fig. 2, where it can be seen that: (*a*) increases

[1] The authors are indebted to Mr. Richard B. McHugh, of the Department of Psychology of Iowa State College, for his helpful advice on the statistical treatment of the data.

[2] This greater variability of the high-height practice curve would seem to support Kimble's notion of inhibition-produced resting responses (12). It has been dealt with in a separate paper (5) which analyzed the prerest practice curves of the 60 *S*s (30 at each height) who received 16 min. of prerest practice. Although the mean high-height curve is significantly more variable than the mean moderate-height curve, the heights do *not* differ in the variability of their underlying individual performance curves. Contrary to theoretical expectations, it appears that the greater variability of the mean high-height curve is due to greater correspondence between high-height *S*s in the points in practice at which they experience performance peaks and depressions.


2 MIN. REST
10 MIN. REST
24 HR. REST
BLOCKS PER MINUTE
ESTIMATED PRACTICE CURVE: Y= 81.0+0.45X
REST AFTER 2 MIN. PRACTICE
REST AFTER 8 MIN. PRACTICE
REST AFTER 16 MIN. PRACTICE
MINUTES OF PRACTICE AT HIGH HEIGHT

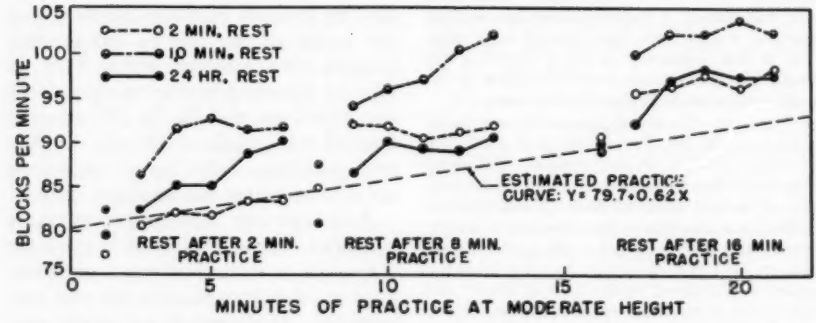

2 MIN. REST
10 MIN. REST
24 HR. REST
BLOCKS PER MINUTE
ESTIMATED PRACTICE CURVE: Y= 79.7+0.62X
REST AFTER 2 MIN. PRACTICE
REST AFTER 8 MIN. PRACTICE
REST AFTER 16 MIN. PRACTICE
MINUTES OF PRACTICE AT MODERATE HEIGHT

FIG. 2. Postrest practice curves at the moderate and high work-surface heights. Plotted points represent the mean performance of ten *S*s. Unconnected points are the groups' performance just prior to rest. The estimated practice curve is from Fig. 1.

in performance following rest appear in almost all groups; (*b*) the initial slope of the postrest curve differs between the groups; (*c*) a number of the curves rise throughout the postrest practice period.

Determination of reminiscence scores.—Reminiscence can be defined as the difference between *S*'s actual postrest performance and his performance if he had not received a rest. In practice, this hypothetical no-rest performance must be estimated from either a control group which does not rest, or by extrapolating *S*'s prerest curve. For the control-group method, *S*'s reminiscence score at a particular minute of postrest practice (M_i) would be defined by the equation:

$$\text{Reminiscence at } M_i = (\text{postrest performance at } M_i) - (\text{performance of control group at } M_i).$$

The disadvantage of this equation is that it does not allow for individual differences in prerest performance, and thus will yield reminiscence scores with considerable error variance.

A reminiscence score from the extrapolation method would be defined by the equation:

$$\text{Reminiscence at } M_i = (\text{postrest performance at } M_i) - (\text{performance from individually-fitted prerest curve extrapolated to } M_i).$$

While this equation does consider *S*'s prerest performance, it suffers from a number of disadvantages: (*a*) the extrapolation may not be accurate, i.e., the prerest trend might not continue if *S* were not given a rest; (*b*) individual curve fitting is laborious; (*c*) individual curve fitting would not be appropriate when only a few prerest points are available, as would be the case for the 60 *S*s of the present experiment who received only 2 min. of prerest practice.

The present study used a modification of the extrapolation method which seemed to remedy most of these disadvantages. Instead of fitting individual curves, extrapolation was based on the mean prerest curve. Thus, linear curves with

the slope of the curves fitted to the mean data of Fig. 1 were passed through S's performance just prior to rest. A reminiscence score for each S was then determined by taking the difference between S's actual postrest performance and the extrapolated linear curve. These operations are reflected in the equation:

Reminiscence at M_i = (postrest performance at M_i) − (performance just prior to rest) − M_i (slope of curve fitted to mean prerest data: .62 for moderate height or .45 for high height).

Scores computed by this equation showed approximately 50% less within-groups or error variance than scores computed by the control-group method. In addition, the curves of Fig. 1 indicate that the extrapolation is reasonably accurate for Ss in the 2- and 8-min. practice conditions, since the prerest curves appear to be essentially linear from 2–16 min. of practice. The experiment does not furnish such a check for Ss in the 16-min. conditions, and thus there is a possibility of extrapolation errors and consequent reminiscence score bias for these Ss.

The differing slopes of the postrest curves of Fig. 2 suggest that the reminiscence scores should be adjusted for warm up (9). For each of the 18 groups a reminiscence–postrest practice function was plotted, and the three consecutive minutes

TABLE 1

BASIC REMINISCENCE DATA OF THE EXPERIMENT

Experimental Condition			Statistic	
Work-Surface Height	Amount Prerest Practice	Length of Rest	Mean Rem. Score*	σ_M
Moderate	2 min.	2 min.	3.45	1.51
Moderate	2	10 min.	8.13	1.10
Moderate	2	24 hr.	5.78	1.49
High	2	2 min.	6.00	2.74
High	2	10 min.	5.73	1.46
High	2	24 hr.	6.16	1.10
Moderate	8	2 min.	5.39	1.26
Moderate	8	10 min.	9.80	2.65
Moderate	8	24 hr.	6.53	2.38
High	8	2 min.	8.23	2.02
High	8	10 min.	9.10	0.94
High	8	24 hr.	5.96	1.80
Moderate	16	2 min.	4.74	0.92
Moderate	16	10 min.	11.12	1.35
Moderate	16	24 hr.	5.70	1.59
High	16	2 min.	13.40	2.41
High	16	10 min.	13.17	2.69
High	16	24 hr.	7.67	2.63

* Reminiscence means are expressed in blocks turned per minute.

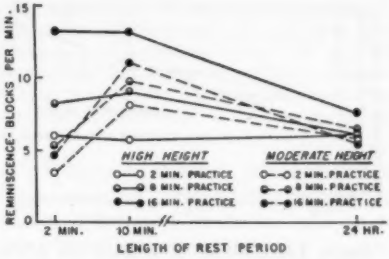


FIG. 3. Reminiscence-rest curves for various amounts of prerest practice at the high and moderate work-surface heights. Each point represents the mean reminiscence score of ten Ss.

of practice which yielded the most reminiscence for the particular group were determined. The S's reminiscence score, adjusted for warm up, was then taken as his average reminiscence score during those 3 min. of postrest practice. Since these scores exhibited heterogeneity of variance which was not removed by a logarithmic transformation, t tests based on Welch's formulae (23) were used in treating the data.

Reminiscence data.—Reminiscence means and variabilities for each of the 18 groups are arrayed in Table 1. The means are plotted in Fig. 3 in the form of reminiscence-rest functions at the various values of prerest practice and work-surface height. Applications of the t test support the conclusions that all reminiscence points are significantly greater than zero, and that the 10-min. reminiscence points tend to be significantly greater than either the 2-min. or the 24-hr. points.[3]

These results are consistent with previous studies (2, 10, 15) which show that, for relatively short rests (10 min. or less), reminiscence increases with rest-period length. Such a result is expected from I_R theory, since the longer rests should allow more complete dissipation of I_R. The decreased reminiscence at long rests

[3] Since the reminiscence means of Fig. 3 are based on independent groups of Ss, a procedure described by Kendall (11, pp. 132–134) may be used to generalize the various rest comparisons over practice and height conditions. Such a procedure yields the following probabilities of the hypothesis that the rest period means are identical: (*a*) For 10-min. rest vs. 2-min. rest: $p < .01$; (*b*) For 10-min. rest vs. 24-hr. rest: $p = .05$; (*c*) For 24-hr. rest vs. 2-min. rest: $p = .58$.

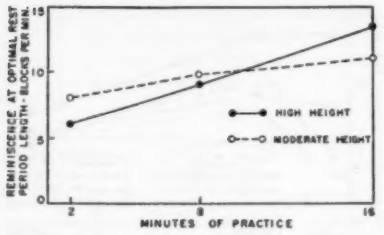


FIG. 4. Reminiscence at optimal rest (peak reminiscence values of curves of Fig. 3) as a function of work-surface height and prerest practice. Each point represents the mean reminiscence score of ten *S*s.

(6 hr. or more) has also been obtained by Ammons (**2**). This decline may be tentatively attributed to forgetting depressing postrest measures at these long rests.

The curves of Fig. 3 also suggest an interaction between height and rest: a 2-min. rest is generally sufficient to generate maximum high-height reminiscence, while moderate-height reminiscence requires at least a 10-min. rest to reach a maximum. In terms of inhibition theory, such results imply that high-height I_R is completely dissipated by shorter rests than moderate-height I_R. Compound t tests (significance of a difference between mean differences) were used to evaluate this aspect of the data by testing the hypothesis that the 2–10 min. rest reminiscence increase is identical for both heights. This hypothesis could be tested with the data from the 2-, 8-, or 16-min. practice conditions. Although none of these three t's were significant, their combined probability (obtained by using Kendall's procedure noted in footnote 3) was .14.

This result is suggestive. If subsequent research indicates that I_R at more effortful conditions dissipates more quickly, a problem would be posed for inhibition theories which assume that greater amounts of I_R require longer rests for complete dissipation (**1**, p. 276). The tendency toward faster dissipation of high-height I_R was allowed for in the critical height comparisons by using points of maximum reminiscence from each reminiscence-rest curve.

These maximum reminiscence values are plotted as a function of height and practice in Fig. 4. It is apparent that the predicted greater high-height reminiscence fails to materialize; only one of the three high-height reminiscence points falls above the comparable moderate-height value. The t tests of Table 2 support the conclusion that reminiscence is comparable at the two work-surface heights, and that the increased reminiscence for longer practice periods is significant.

The reminiscence-practice functions of Fig. 4 do not exhibit the decrease in reminiscence after approximately 10-min. practice which characterizes previous results (**2, 10, 12**) and which has been attributed to losses in motivation late in practice (**12**). A pos-

TABLE 2

t COMPARISONS OF THE REMINISCENCE MEANS OF FIG. 4

Comparison	Mean Diff.	$\sigma_{\text{Mean Diff.}}$	df[a]	t[b]
High-height minus moderate-height reminiscence				
2-min. practice	−1.97	1.56	18	−1.26
8	−0.70	2.82	12	−0.25
16	2.28	2.76	15	0.83
Reminiscence at 16 min. minus reminiscence at 2-min. practice				
Moderate height	2.99	1.75	18	1.71
High height	7.24	2.66	18	2.73*

[a] Calculated by Welch's formulae (**23**).
[b] One-tailed tests (**17**) used for testing high-height minus moderate-height reminiscence; two-tailed tests used for testing 16-min. minus 2-min. reminiscence. In the latter case, Kendall's procedure (**11**, 132–134) yields a probability of 0.02 for the null hypothesis for both heights combined.
* t significant at the 5% level of confidence, i.e., greater than 2.10.

sible explanation of the continued rise of our reminiscence-practice function is that the frequent knowledge-of-results incentive maintained *S*'s motivational level throughout the practice period.

Discussion

Reminiscence in the block-turning task has been shown to be directly related to amount of prerest practice and length of rest interval. In general, these results confirm previous findings for various kinds of motor tasks and are consistent with expectations from I_R theory. However, in neither prerest performance nor reminiscence has work-surface height been a differentiating variable. These findings are not consistent with theoretical expectations. It will be recalled that I_R theory predicts a proportionality between task effortfulness and I_R. Following from this is the expectation that performance should be poorer and reminiscence should be greater for the high height than for the moderate height.

The failure to find differences in prerest performance for the two heights also contradicts the previous study (6) in which high-height performance was found to be poorer than moderate-height performance. There are several methodological differences between the two studies which may account for this contradiction. (*a*) The previous study utilized a latin-square design in which each *S* served at several heights. This might allow for certain height sets to develop which could depress high-height performance. (*b*) The two studies differ in factors influencing motivation. In the previous study knowledge of results was given after each 3-min. work period while in the present study such knowledge was given after each 1-min. work period. Also, *Ss* in the previous study were fully informed of the purpose of the experiment and the conditions under which they were to work. The *Ss* in the present study were in a much more demanding situation. Their only information was that they were required to work for some indefinite period of time at a task which apparently was so stressful that *E* gave them an initial opportunity to withdraw from the experiment. (*c*) In the previous study 3 min. of work was followed by 2 min. of rest while, in the present study, work was continuous. Assumptions about the rate of development and dissipation of I_R as a function of height could be made which would account for performance differences in the previous study and not in the present one.

Without additional research there is no way of determining which of these differences between the two experiments are critical and consequently what interpretation should be placed on the height variable. In the absence of such research, we can point up alternative interpretations of the height variable, and then note the theoretical implications of the present study's results under these interpretations.

Alternative interpretations of the height variable.—The interpretation that height does not influence either effort or inhibition is rejected at the outset. It is our presumption from observing our *Ss* that the high-height task was more difficult and more effortful than the moderate-height task. Profanity, requests to stop work, agonized facial expressions, and hostile verbal comments were frequently observed for high-height *Ss* and rarely for moderate-height *Ss*.

A second interpretation is that height influences effortfulness but not inhibition. This merits consideration. The antagonistic behavior of our high-height *Ss* would be attributed to effortfulness, while the failure of the high height to depress performance and enhance reminiscence would be consistent with the notion of equal inhibition. This is a parsimonious interpretation since it rests on the principle that differential amounts of I_R can only be inferred from acquisition and retention data. The difficulty with the interpretation is the discontinuity between the effortful and inhibitional effects of height which it assumes. However, there may be a

good reason for this apparently awkward discontinuity. In the previous study, *Ss* localized high-height muscular strain primarily in the back and upper arms (**6**, p. 293). Thus high-height effortfulness could be localized in muscle systems so distant from the wrist and finger muscle systems involved in the block-turning response that its inhibitional effects on the response are negligible.

A final possibility is that height influences both effort and inhibition. Such an interpretation requires an explanation of why neither performance nor reminiscence was influenced by height in the present study. There are two factors which could be invoked: the high height may have had greater habit strength or greater motivation associated with it in the present study. Either of these factors could have offset the greater high-height inhibition.

Theoretical implications of results under alternative interpretations of the height variable.—If it is assumed that the height variable influences effort, but not inhibition, the results of the present study would not be considered critical for inhibition theory. To our knowledge, there is no case in the literature on motor learning in which equal acquisition curves were followed by unequal amounts of reminiscence. Therefore, having shown that work-surface height differences did not appreciably influence acquisition, we might not expect differences in reminiscence for the two heights.

If, on the other hand, it is assumed that the height variable does influence both effort and inhibition, then the results of the present study could be viewed as inconsistent with inhibition theory. This would be true irrespective of whether habit strength or motivation is advanced as the factor which concealed the high height's greater inhibition prior to rest. If habit strength is used, then, according to theory, the inhibition should dissipate over rest leaving the high height with superior postrest performance and consequently greater reminiscence. A similar argument can be advanced for motivation. Current inhibition theory assumes (**13**, p. 500), with some empirical basis (**14**, **22**), that more highly motivated *Ss* will tolerate more I_R and hence show more reminiscence. If greater high-height motivation is used to explain the prerest results, it follows that greater high-height reminiscence would be the theoretical expectation.

This contradiction could be removed by invoking conditioned inhibition (a permanent form of inhibition; or, a learned resting response as described by Kimble (**12**)). If it is assumed that conditioned inhibition is greater at the high height than at the moderate height, the failure to find greater reminiscence for the high-height group could be handled. We are not particularly impressed by this explanation because the usefulness of the conditioned inhibition construct has yet to be adequately demonstrated.

In summary, then, we conclude that we have found no evidence for a proportionality between effortfulness of task and amount of inhibition. This finding appears to pose difficulties for inhibition theory, although it is possible that our method of varying effortfulness is inappropriate. However, the finding seems to have some generality in view of Eckstrand's (**4**) and Helmick's (**7**) comparable results with the pursuit rotor.

Summary

An experiment was performed to test the applicability of the Hullian reactive inhibition construct to performance decrements previously ob-

tained in a block-turning task at high work-surface heights. To test the prediction of greater high- than moderate-height reminiscence, pre- and postrest performance measures were obtained from 180 *Ss* working under various conditions of amount of prerest practice (2, 8, 16 min.) and rest-period length (2 min., 10 min., 24 hr.) at the two heights. The major results:

1. The high and moderate heights were equivalent in prerest performance.
2. The high and moderate heights were equivalent in reminiscence.
3. For both heights, reminiscence increased with increases in amount of prerest practice.
4. For both heights, reminiscence was markedly decreased at the long (24 hr.) rest.

While the last two results were consistent with inhibition theory and previous research, the first two results could not be handled readily within such a framework.

(Received May 9, 1952)

References

1. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
2. Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.*, 1947, **37**, 393–411.
3. Ammons, R. B. Acquisition of motor skill: III. Effects of initially distributed practice on rotary pursuit performance. *J. exp. Psychol.*, 1950, **40**, 777–787.
4. Eckstrand, G. A. Post-rest performance in motor learning as a function of amount of work performed in pre-rest practice. Unpublished Ph.D. dissertation, Northwestern Univer., 1949.
5. Ellis, D. S. Periodicities in motor performance as a function of work-surface height. *Trans. Iowa Acad. Sci.*, 1951, **58**, 331–339.
6. Ellis, D. S. Speed of manipulative performance as a function of work-surface height. *J. appl. Psychol.*, 1951, **35**, 289–296.
7. Helmick, J. S. Pursuit learning as affected by size of target and speed of rotation. *J. exp. Psychol.*, 1951, **41**, 126–138.
8. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
9. Irion, A. L. The relation of 'set' to retention. *Psychol. Rev.*, 1948, **55**, 336–341.
10. Irion, A. L. Reminiscence in pursuit-rotor learning as function of rest and amount of pre-rest practice. *J. exp. Psychol.*, 1949, **39**, 492–499.
11. Kendall, M. G. *The advanced theory of statistics.* Vol. II. London: Griffin, 1948.
12. Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.
13. Kimble, G. A. Performance and reminiscence in motor learning as a function of degree of distribution of practice. *J. exp. Psychol.*, 1949, **39**, 500–510.
14. Kimble, G. A. Evidence for the role of motivation in determining the amount of reminiscence in pursuit rotor learning. *J. exp. Psychol.*, 1950, **40**, 248–253.
15. Kimble, G. A., & Horenstein, B. R. Reminiscence in motor learning as a function of length of interpolated rest. *J. exp. Psychol.*, 1948, **38**, 239–244.
16. Krueger, W. C. F. The influence of amount limits and time limits upon the rate of work. *J. appl. Psychol.*, 1937, **21**, 113–118.
17. Marks, M. R. Two kinds of experiment distinguished in terms of statistical operations. *Psychol. Rev.*, 1951, **58**, 179–184.
18. Montgomery, K. C. An experimental investigation of reactive and conditioned inhibition. *J. exp. Psychol.*, 1951, **41**, 39–51.
19. Ross, S., & Bricker, P. D. The effect of an amount-set on a repetitive motor task. *J. exp. Psychol.*, 1951, **42**, 39–43.
20. Siegel, P. S. Reactive inhibition as a function of the number of response evocations. *J. exp. Psychol.*, 1950, **40**, 604–608.
21. Thompson, M. E. Reactive inhibition as a factor in maze learning: I. The work variable. *J. exp. Psychol.*, 1950, **40**, 131–134.
22. Wasserman, H. N. The effect of motivation and amount of pre-rest practice upon inhibitory potential in motor learning. *J. exp. Psychol.*, 1951, **42**, 162–172.
23. Welch, B. L. The generalization of 'Students' problem when several different population variances are involved. *Biometrika*, 1947, **34**, 28–35.

FAILURES OF ATTENTION IN SELECTIVE LISTENING

D. E. BROADBENT[1]

Medical Research Council, Applied Psychology Research Unit, Cambridge, England

In a previous paper (3), it was shown that identification of the correct voice to listen to, in a babel, was sometimes more difficult than listening to it once it was known. Three points of interest arose from this work: first, that continuously listening to one voice appeared in those results to be no easier than changing from voice to voice in obedience to a visual cue. Yet it was felt that repeated alterations in the complex stimulus to be selected should afford more opportunities for failure in correct selection. Second, even when the correct voice was known there still occurred many failures to answer it correctly. It was possible that these failures were simply due to masking, and would have occurred even if the other voice had been a meaningless noise: other workers have given evidence to show that the meaning of irrelevant masking speech does not increase its effect on intelligibility (6). Yet this seemed unlikely from introspection. Third, it was of interest that two simultaneous streams of information could not be stored until it became evident which was relevant. Is this difficulty dependent on the simultaneity of the sounds at the ear, or rather on a need for sequences of related items to be kept separate perceptually?

The following experiments help to clarify all these points, and are linked by a common feature, namely, that even when there is a nonauditory indication of the voice to be listened to, mistakes may appear which can be ascribed to failure to select the relevant from amongst the irrelevant information (failures of attention).

[1] This work was carried out while the author was in receipt of a grant from the British Medical Research Council, and thanks are due to the Royal Navy both for supplying *S*s and for providing technical equipment.

Procedure

The general procedure of the two previous papers was adopted (2, 3); that is, *S* was asked a series of Yes-No questions from the loudspeaker of a tape recorder about a visual display. The recorder was the Ferrograph Type C used in the latter report, and the minor details were the same. The form of the questions was altered in this case, however, to ensure that every word was essential to the meaning. To bring this about each of the five numbered sections of the visual display carried two conventional objects, and a question consisted simply of the names of two objects and the number of a section: for example, "Square Cross 3." This was interpreted as "Is there a Square and a Cross on Section 3?" The *S* answered "Yes" or "No" accordingly. The questioner gave his own name (GDO or Turret) before each of the questions, and *S* put the same name before his answer. Thus the full dialogue might be "GDO Heart Circle 2," with the answer, "GDO No." All *S*s were trained to proficiency in this routine on the day before the experiment proper.

Group I (20 *S*s) received two sequences of questions of this type, with the complication that two messages always started synchronously, one from each voice. The part of GDO was always read by one voice and that of Turret by the other. In one sequence *S* was told always to answer one voice (GDO for half the *S*s and Turret for the other half), while in the other sequence the names of the two questioners formed part of the visual display and *E* indicated before each question which voice was to be answered. The situation differed from that of Group II of the last paper (3), in that the sequences of questions used were much longer, containing 48 pairs of simultaneous questions instead of 12.

It should be noted that in all other groups of the present paper a sequence consisted of 24 pairs of questions.

Group II (14 *S*s) received two sequences in which they were required to listen to two questions completely before answering them both. In one sequence one voice asked a question, then the other asked one, and then there was a gap

TABLE 1

LAYOUT OF MESSAGES OF DIFFERENT TYPES

Voice I Voice II Subject	GDO Heart Cross One Turret Circle Square Two GDO Yes	Type X
Voice I Voice II Subject	GDO Heart Cross One Turret Circle Square Two GDO Yes Turret No	Type A
Voice I Voice II Subject	GDO Heart Cross One Turret Circle Square Two GDO Yes Turret No	Type B

for the two answers before the cycle was repeated (Type A in Table 1). In the other sequence alternate words were heard from the two voices, until both questions were complete and then a gap was left for the two answers before the cycle was repeated (Type B in Table 1). The rate of speech was approximately the same in both sequences, if anything, the Type A one being faster.

Group III (12 *Ss*) heard the same questions as Group II, but were told only to answer one voice. In all cases this was GDO, who was slightly the less intelligible.

Group IV (10 *Ss*) received as one sequence of questions the Type B (alternate word) sequence of Group II, with instructions only to answer GDO. As another sequence, they received a series of Type B questions in which the parts both of GDO and Turret were spoken by the same voice (GDO in the other series). Care was taken to ensure similar timing in the two recordings. Again *Ss* were instructed only to answer the questions having the call-sign GDO.

Group V (12 *Ss*) received the Type B (alternate word) sequence of Group II with instructions only to answer one voice (GDO for half the group and Turret for the other half). They also received a similar series in which the two voices changed roles in random fashion from question to question, each being sometimes GDO and sometimes Turret. Each question, however, was spoken by one voice throughout. Again *Ss* were instructed only to answer questions from one source, those who answered GDO in the other sequence having to answer GDO in this one, and those who answered Turret having to answer Turret.

In every group the order of conditions was reversed on alternate *Ss*. Groups II and III were composed of *Ss* who had served in Group I 24 hours previously: Group I is smaller than the sum of these groups because the results for 6 *Ss* were discarded owing to failure of the response recording system (unknown to *S*). Otherwise all groups were of different individuals. All *Ss* were naval ratings aged between 19 and 30, and as in the two previous reports, the degrees of difference between the different conditions did not appear to be related to their Service trades.

RESULTS

Maintaining and shifting attention.—The main results for each group are presented in Table 2. In Group I the condition for listening steadily to one voice was compared with that of changing repeatedly from voice to voice in obedience to a visual signal. In the former case half the group listened to GDO and half to Turret. Considering only the first half of each sequence, we find no significant difference between the two conditions ($p > .05$). There is, however, a significant practice effect (not shown in Table 2), ($.01 < p < .05$). These results agree with those from the shorter sequences used in the last report. When we consider the second half of each sequence, however, the practice effect disappears ($p > .05$) and *the condition of listening steadily to one voice is quite clearly better than that of shifting* ($p < .001$).

A possible reason for the absence of this difference at low levels of practice is that a voice must be heard for more than a dozen or so messages before its characteristics are sufficiently well

TABLE 2

CONDITIONS GIVEN TO THE DIFFERENT GROUPS

Group	*n*	Conditions		Percentage of Mistakes		Number of *Ss* Performing Better	
		a	*b*	*a*	*b*	*a*	*b*
I	20	Type X messages; listen to one voice throughout series.	Type X messages; listen to one voice in each message, but change from voice to voice between messages in obedience to visual cue.	Early in run: 14 Late in run: 9	16 19	10 16	6 2
II	14	Type A: answer both voices.	Type B: answer both voices.	31	63	12	2
III	12	Type A: answer only GDO.	Type B: answer only GDO.	0	10	8	0
IV	10	Type B: answer only GDO: voice of Turret differs from that of GDO.	Type B: answer only GDO: voice of Turret the same as that of GDO.	16	42	8	2
V	6	Type B: answer only GDO: voice of Turret differs from that of GDO; same voice always has same call sign.	Type B: answer only GDO: within each message voice of Turret different to that of GDO, but voices change roles between messages	12	14	6	5
	6	As above, but answer only Turret.	As above, but answer only Turret.				

The variance being highly unstable in these results, tau ranking techniques were used to assess significance, and the number of *Ss* performing better in each condition is given above as a guide to the reader. Where Column 7 + Column 8 < *n*, the remaining *Ss* showed no difference between conditions.

learned for it to be positively identified immediately after a pause. Consequently in a series of messages from a strange voice identification must be re-established in every message and the situation is essentially similar to that of shifting from voice to voice. Further evidence on this point appears in the results from Group V.

Mixed and separated messages.—In Group II there was no actual overlapping of the sounds being listened to, in so far as words from two voices did not occur at the same time. But in one condition (Type A) the sequence of related sounds forming each message was continuous in time, while in the other (Type B) each word was separated from the next by a word from another message. Although in both cases no answer was to be given until both messages had been completed, *it is far easier to answer two messages which occupy different periods of time than to answer two which occupy scattered portions of the same period* ($p < .001$). A mistake in Table 2 is a case in which at least one of the two messages was not answered correctly. It will be noted that on the Type B questions only a few more than the chance score of one in four are correct. Consequently it seems that the extreme difficulty of listening to two voices at once (3) need not be due to the actual simultaneity of the sounds, but would arise in any case from the intermingling of the messages.

Irrelevant speech inside and outside the sequence.—Group III received the same two conditions as Group II, except that everything said by "Turret"

was to be ignored, and only that question, in each pair, which was said by "GDO" was to be answered. It is hardly surprising that in each condition these questions were much more often answered correctly by Group III than by Group II: for each type $p < .0001$.

Less obvious is the finding that even when one voice is to be neglected, Type A is easier than Type B ($p < .01$), that is, *irrelevant speech produces a definite effect when it occurs between the sections of a relevant message*, as compared with the effect of presenting such speech before or after the relevant message. It seems likely, then, that some of the failures in listening to one message in the presence of another irrelevant one are not due to masking, since they still occur even when there is no overlapping of sound.

The cues used in discarding irrelevant information.—In Group IV we are comparing the condition of listening to messages in one voice while ignoring intruding words from another voice, with the condition of listening to messages while ignoring intruding words in the same voice. The latter is far harder than the former ($p < .01$). In Group V, on the other hand, we compare two conditions of listening to the message following a particular call-sign, while ignoring intruding words in another voice. In one condition, the call-sign always belongs to a particular voice, while in the other condition, the voice using the call-sign changes randomly from question to question. There is no evidence in the present results that these two conditions differ at this low level of practice ($p > .05$). This supports the suggestion made earlier that when an unfamiliar voice is speaking a series of messages, it must be identified by the listener at the beginning of each message. At the same time it is clear from the results in Group IV that discarding of irrelevant information does not take place solely on the basis of the time relations between different sounds; voice characteristics form a valuable additional cue.

Discussion

These results have various practical implications, such as the reduction of speech chatter in control rooms even though it is not actually masking the desired signal. More generally, they confirm that the process of selecting part of the presented sounds for analysis and response (which was shown to exist in (3), is not infallible and may lead to errors distinguishable from those due to masking. On the one hand difficulties which appear in changing the voice to be selected do not occur when one voice is always to be selected, and so cannot be due to simple masking; and on the other hand irrelevant information increases errors, presumably because it is not discarded, even though it is not simultaneous with the relevant information. It is likely that such errors will appear only when some fairly similar irrelevant stimulation is present to render selection difficult; the data from Group IV show that increasing the similarity decreases the efficiency of selection.

It will be recalled that it was pointed out (2) that classical work on "distraction" usually presented some irrelevant stimulus both with the stimulus for the task concerned and during the process of organizing responses to the task. The paper referred to separated the latter of these two conditions and showed that a significant effect could only be found when the "distracting" stimulus had some importance for *S*. Irrelevant messages showed no effect on simultaneous speech. In the present work, however, we have been considering

overlapping of stimulation rather than overlapping between stimulus and response, and the effects are far more serious. Even irrelevant messages produce an effect although they are not strictly simultaneous (Group III). It seems probable that the confusing character of existing results on "distraction" is partly due to inadequate separation of the two conditions dealt with earlier (**2**) and in the present paper.

One of the statements common in older work is that attention can only be given to one task at once, but that two or more tasks can be performed by rapid alternation of attention (**8**, p. 708 f.). The results from Group II place a serious limitation on this statement. They also pose the additional problem of understanding why sequences of related stimuli cannot be allowed to interweave in this manner. It is worth reminding the reader of the fact that the number of possible messages formed by a sequence of n signals is far greater than n times the number possible for one signal (**7**): this is, of course, why information is conventionally defined as a logarithmic function. Much of the meaning of a spoken word is not contained in the word itself but in its relation to preceding and succeeding words. If, then, there is a limit to the informational capacity of the perceptual mechanism, and fresh stimulation must be neglected when this capacity is already being utilized (as the previous work strongly suggests), it is reasonable that the limit should be reached by the message as a whole and not by individual words. Extra words occurring between the words of a sequence find the listener already dealing with nearly as much information as he does at the actual instant of hearing a word.

Indeed, speech is only a special case of a principle which may apply to all tasks. A stimulus of any kind often contributes most information by its position in a stream of stimuli and not by its isolated occurrence (**1**). Thus, if a person is required to deal with two streams of signals there is an effect on his performance which is not the same as that of doubling the rate of arrival of signals in a single stream (**4, 5**). The particular case of speech is interesting, however, because it excludes such factors as eye movements.

Summary

The present paper contains a series of experiments on the selection out of a mass of speech of parts which are relevant. Group I found it easier to listen steadily to one voice in a mixture rather than change from voice to voice in obedience to a visual cue. This applied only when the voices were familiar, so that one may suspect that an unfamiliar voice has to be reidentified every time it speaks.

Group II was required to listen to two questions and then answer them both before the cycle was repeated. In one case, the questions followed one another in time, while in another, alternate words were heard from each voice. The latter condition was far harder, showing that the difficulty of listening to two messages at once is not purely a matter of difficulty in hearing (masking).

Group III met the same conditions as those in Group II, but were only expected to answer one of the two questions. The alternate word condition was still the harder, so that neglect even of irrelevant information is not perfect in this case.

Two groups of *Ss* underwent the conditions of Group III with variations in the vocal cues available for discarding the irrelevant information. It appeared that if the same voice read both relevant and irrelevant words, it was harder to answer the

relevant ones than if different voices read the two types of information. On the other hand, when the voice using the call-sign to be answered varied randomly from question to question, performance was as good as when the same voice used the call-sign throughout the session. This also seems to indicate that reidentification occurs in every message with an unfamiliar voice.

These results have certain implications for the arrangement of communications centers; they are also related both to earlier work on distraction and to modern studies of load stress.

(Received May 26, 1952)

References

1. Bartlett, F. C. *Remembering.* Cambridge: Cambridge Univer. Press, 1932.
2. Broadbent, D. E. Speaking and listening simultaneously. *J. exp. Psychol.*, 1952, **43**, 267–273.
3. Broadbent, D. E. Listening to one of two synchronous messages. *J. exp. Psychol.* 1952, **44**, 51–55.
4. Conrad, R. Preliminary factory studies of the effect of speed and load stress on cotton winding. Cambridge: Appl. Psychol. Unit. Report No. 147, 1951.
5. Conrad, R. Speed and load stress in a sensori-motor skill. *Brit. J. industr. Med.*, 1951, **8**, 1–7.
6. Miller, G. A. The masking of speech. *Psychol. Bull.*, 1947, **44**, 105–129.
7. Miller, G. A. *Language and communication.* New York: McGraw-Hill, 1951.
8. Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1938.

SITUATIONAL INSTRUCTIONS AND TASK ORDER IN RECALL FOR COMPLETED AND INTERRUPTED TASKS[1]

DAVID G. HAYS

Harvard University

Zeigarnik discovered in 1927 that, in general, interrupted tasks are recalled more frequently than completed tasks (**9**). She also found that upon interpolation of an emotional situation between task performance and recall, completed tasks were recalled more frequently than interrupted tasks. Later work with interrupted tasks has tended to follow one of three or more lines. (*a*) Following Zeigarnik, researchers of the Lewinian school have performed experiments and written reviews which applied their explanatory concepts to the area (e.g., **3, 4**). (*b*) Experimental psychologists have sought to make more precise specifications of the conditions relating to the "Zeigarnik effect" (e.g., **2, 5, 6**). (*c*) More recently, "personality" variables have been applied by workers of other schools than Lewin's (e.g., **1, 7, 8**).

Two experimental variables are examined in the present work. Rosenzweig (**7**) and others have shown that such variables as "involvement" have an influence on recall for complete and incomplete tasks. Involvement has regularly been controlled by situational instructions; the contrast has frequently been between instructions that the subjects are taking an "intelligence test" on the one hand, and that they are helping to "standardize materials" on the other. These two situations were used in the experiment reported here.

[1] Prepared under the direction of James Olds, Department of Social Relations, Harvard University.

Zeigarnik (**9**) showed that interpolation of an emotional situation between task performance and recall has an effect on differential recall for complete and incomplete tasks. Many other experimenters have found effects on recall associated with serial order of tasks. The second variable employed was the nature of the task following interruption or completion.

METHOD

Two forms of *situational instructions* were used. In Situation A, *Ss* were instructed that the tasks they were to perform had been standardized as a measure of intelligence. They were told that their scores would be compared with their scores on other intelligence tests which they had taken; hence they were required to sign their papers. Finally, *Ss* were instructed that they should be able to complete all of the tasks which would be presented in the time allowed.

In Situation B, *Ss* were instructed that they were being asked to help standardize some tasks which might prove useful as a measure of intelligence. They were told *not* to sign their papers. The fact that they might not complete some tasks was stressed.

The tasks were of two forms, presented alternately. *Experimental* tasks were completed in some cases and interrupted in others according to a prearranged schedule. *Interpolated* tasks, following all experimental tasks, were in every case completed.

Two forms of interpolated tasks were prepared. Interpolated tasks of Type I comprised a set of passages from *War and Peace;* interpolated tasks of Type II consisted of a set of passages written expressly for the experiment, styled to be as dull as possible. At the conclusion of each interpolated task, *Ss* rated the task on a scale of "interest" in unit steps from 1 (low interest) to 5 (high interest). Differentiation between the tasks is indicated by the means of these ratings. The mean for all tasks of Type I was 3.41 ($s_{\bar{x}}$ = .0414); that for all tasks of Type II was 2.75 ($s_{\bar{x}}$ = .0156).

The experiment was designed for group presentation. In order to assure completion of some experimental tasks, and interruption of others, with time allowed for each task held constant, two forms of experimental task instructions were prepared. All experimental tasks consisted of lists of qualities and lists of objects; the *Ss* were instructed to locate the objects having the required qualities. The two types of task instructions were applied to identical lists of qualities and objects. Instructions of Type 1 called for selection of the number of objects on the list which actually possessed the required qualities; instructions of Type 2 called for selection of one more object than the number actually included in the list. Hence, Type 1 instructions allowed the tasks to be completed, while Type 2 instructions led perforce to incompletion. Some *Ss* became aware of the fact that Type 2 instructions were not legitimate, however, and others gave incorrect solutions. These *Ss* were not included in the analysis.

Following situational instructions, four experimental tasks were presented to each *S*. Thirty seconds were allowed for each experimental task. An interpolated task was presented after each experimental task. Three minutes were allowed for reading each interpolated-task passage.

Each *S* heard only one kind of situational instructions (Situation A or B); each *S* performed two experimental tasks of Type 1 and two tasks of Type 2 in alternating order (1, 2, 1, 2 for half the *Ss*; 2, 1, 2, 1 for the remaining *Ss*). Each *S* was presented in all with four interpolated tasks, but all of the same Type (I or II).

Recall was measured following the fourth interpolated task. Three "qualities" were listed for each task. The *S* was asked to recall these four lists. Recall scores *for tasks* thus ranged from one to three, with four observations for each *S*.

The subjects were 27 students in several courses in social relations at Harvard University.

Results

Mean recall scores for all conditions, with their standard errors, are given in Table 1. In order to perform an analysis of variance, the following steps were carried out. First, the mean of the two recall scores observed for each *S* with all conditions constant was taken. Second, the mean recall score obtained with task instructions of Type 2 (which led to task interruption) was subtracted from the mean recall score of the same *S* obtained with task instructions of Type 1 (which led to task completion). Thus *positive values* of these difference scores show the advantage of *completed* tasks in recall, while *negative values* show the recall advantage of *interrupted* tasks. Finally, in order to have proportional numbers of observations in all cells, two *Ss* were discarded by means of a random number-selection process. The mean recall difference scores for all conditions, with their standard errors, are given in Table 2 for the 25 *Ss* included in the final analysis of variance.

TABLE 1

Recall Scores for All Task Conditions

Instructions		Interp. Task	Recall Score		N
Situation	Exp. Task		Mean	σ_M	
A	1	I	2.79	.11	14
		II	2.72	.18	18
	2	I	2.43	.17	14
		II	2.78	.10	18
B	1	I	2.63	.07	8
		II	2.29	.09	14
	2	I	2.38	.14	8
		II	2.64	.02	14

The analysis of variance performed on mean difference scores is presented in Table 3.

The analysis of variance strongly supports the hypothesis that the form of the interpolated task had an effect on recall for the preceding task. The

TABLE 2

Recall Difference Scores for All Task Conditions

Instructions (Situation)	Interpolated Task	Mean	σ_M	N
A	I	+.33	.17	6
	II	−.06	.22	9
B	I	+.25	.37	4
	II	−.33	.25	6

TABLE 3

ANALYSIS OF VARIANCE OF MEAN DIFFERENCE SCORES

Source	*df*	Mean Square	*F*	*p*
Interpolated Task	1	1.307	17.197	.005
Situation Instruction	1	.240	3.158	.08
Interaction	1	.054		
Within Subclasses	21	.076		

value of *F* found was significant well beyond the .005 probability level. The difference in situational instructions was found to be significant beyond the .10 level, but not at the .05 level. Interaction is not significant.

DISCUSSION

The general form of this experiment may be compared with that of the usual study of the effects of completion and interruption on recall. It will be noted first that most *E*s have presented tasks of the kind designated here as "experimental" in succession, without interpolation. No control other than randomization of task order has regularly been applied to the variable measured in this experiment by *S*s' "interest" ratings of the interpolated tasks. If all of an *E*'s tasks are of Type I, he should, according to the present results, find completed tasks recalled more often than interrupted tasks. If all of his tasks are of Type II, on the other hand, he would be expected to find recall advantage for interrupted tasks. If the two types of tasks are combined in one experiment, results in either direction, or results showing no advantage for either condition, might be found.

Rosenzweig's work has pointed more clearly to the importance of situational instructions. The present experiment serves only to underline the the fact that relative recall for completed and interrupted tasks is affected by the situation in which the tasks are performed.

It is to be concluded that the effect of interruption on recall varies with other factors. Among these factors are the situation defined by *E* in his instructions and the *nature of the task following interruption or completion.* The advantage in recall of completed tasks seems to be maximized in the present experiment when the situation is defined as a test of intelligence (which Rosenzweig calls the involvement situation) and when the following task is of most interest to *S* as measured by his own rating. The recall advantage of interrupted tasks is maximized here when the situation is defined as experimental and when the following task is of least interest.

SUMMARY

Twenty-seven *S*s participated in an experiment on recall for complete and interrupted tasks. The situational instructions and the nature of the task following completion or interruption were varied.

Sixteen of the *S*s were instructed that an intelligence test would be administered to them (Situation A). Eleven were instructed that they were to help standardize test materials (Situation B). Each *S* performed four experimental tasks, of which two were completed and two interrupted. Each experimental task was followed by an interpolated task consisting of a passage of complex (Type I) or simple (Type II) prose.

Completed tasks are found to be recalled more frequently when the following interpolated task is complex, while interrupted tasks are more frequently recalled preceding simple tasks. Completed tasks are recalled more frequently when *S* is told that he is assisting *E*. The contrary holds for interrupted tasks. No interaction

is noted between instructions and interpolated task.

These results indicate that the effect of interruption on task recall varies with other factors, including the nature of the task following interruption or completion, and the general instructions which *E* presents. It is suggested that these variables may account for some previously unexplained variability in the results of similar experiments.

(Received May 26, 1952)

References

1. Alper, T. G. Memory for completed and incompleted tasks as a function of personality: an analysis of group data. *J. abnorm. soc. Psychol.*, 1946, **41**, 403–420.
2. Boguslavsky, G. W. Interruption and learning. *Psychol. Rev.*, 1951, **58**, 248–255.
3. Koffka, K. *Principles of Gestalt psychology*. London: Kegan, Paul, Trench, Trubner, 1935.
4. Ovsiankina, M. Die Wiederaufnahme unterbrochener Handlungen. *Psychol. Forsch.*, 1928, **11**, 302–379.
5. Pachauri, A. R. A study of Gestalt problems in completed and interrupted tasks. Part I. *Brit. J. Psychol.*, 1935, **25**, 365–381.
6. Pachauri, A. R. A study of Gestalt problems in completed and interrupted tasks. Part II. *Brit. J. Psychol.*, 1935, **25**, 447–457.
7. Rosenzweig, S. Preferences in the repetition of successful and unsuccessful activities as a function of age and personality. *J. genet. Psychol.*, 1933, **42**, 423–440.
8. Rosenzweig, S. An experimental study of repression with special reference to need-persistive and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, **32**, 64–74.
9. Zeigarnik, B. Über das Behalten von erledigten und unerledigten Handlungen. *Psychol. Forsch.*, 1927, **9**, 1–85.

THE EFFECT OF LONG INTERTRIAL INTERVALS ON THE LIMIT OF EYELID CONDITIONING [1,2]

MARTIN R. BARON [3]

State University of Iowa

Available experimental evidence has been quite consistent in indicating that relative massing of trials in classical conditioning decreases the degree of learning exhibited (**2, 3, 4, 6, 7, 8**). However, it is not possible to determine from these data the exact empirical nature of the function relating the two variables.

In most of the investigations reported (**2, 4, 6, 7**), the relative effects of *only two* intertrial intervals were compared. Unfortunately, differences in experimental method prohibit integration of the results of any two *E*s. There are only two studies (**3, 8**) in which the relative effects of *more than two* intertrial intervals are compared. Presenting the stimuli 3, 9, and 18 times per minute, Calvin (**3**) obtained little or no eyelid conditioning in the two relatively massed groups; thus his data are not sufficient to determine the over-all functional relationship between the variable and the level of performance. Spence and Norris (**8**) conditioned the eyelid response in four groups of *S*s with intertrial intervals of 9, 15, 30, and 60 sec. each. They interpret their data to ". . . suggest that the amount of conditioning as revealed by frequency measures is a negatively accelerated increasing function of the time between trials" (**8**, p. 719).

The present investigation, conducted concurrently with the Spence and Norris study (**8**), attempts to test one aspect of the hypothesis that the function relating the intertrial interval and the frequency measure of conditioning is negatively accelerated.[4] More specifically, assuming that this relationship approaches a limit, the purpose of this study is to determine, if possible, two widely separated intervals at which no differential response frequency would be obtained. In view of tentative findings in a preliminary study that there is a difference between levels of performance with intertrial intervals of 90 and 120 sec., the intervals chosen in this study were 120 and 270 sec. In order to keep both groups from reaching 100% frequency of response, it was not possible to utilize the most commonly employed conditioning variables. Instead of these, a trace conditioning procedure was used with an interval of 1.50 sec. between the CS and UCS.

METHOD

Apparatus and stimuli.—The apparatus used in the present study was the Dodge photochronograph which has been described elsewhere (**1**). Except for the particular eyelash and the motion pictures employed here, the apparatus is the one

[1] This paper is a portion of a thesis submitted to the department of psychology of the State University of Iowa in partial fulfillment of the requirements for the degree of Master of Arts. The writer is indebted to Dr. Kenneth W. Spence for his guidance in the formulation of the problem, throughout the experimental investigation, and in preparation of the manuscript.

[2] A summary of this paper was read before the Midwestern Psychological Association meeting in Chicago, 1948.

[3] Now at Kent State University.

[4] It is possible to extend Reynolds' theory concerning rate of presentation of the CS (**6**) to predict the relationship between the intertrial interval and the limit of conditioning. The function predicted is similar to a negatively accelerated growth curve. As pointed out by Spence and Norris (**8**), essentially the same prediction can be made from a work-decrement hypothesis.

used by Reynolds (6) and Kimble (5). A paper eyelash, ¼ in. wide, and 2½ in. long, attached to the center of the upper eyelid of the left eye by Scotch Tape, interrupted a light beam and cast a shadow on "Kodagraph Standard Fast Projection" photographic film. The wash light source was a Bausch and Lomb 35-mm. projector.

The *S*s sat in a straight-backed chair with an adjustable seat and headrest. A biting board was employed to insure that each *S*'s head would be in the same position from trial to trial.

The CS was a .10-sec., 1000-cycle tone, 15 db's above the threshold, produced by a beat frequency oscillator and delivered to *S* through a pair of earphones. Coincident with the onset of the tone was a click made by an electromagnetic signal marker. The UCS, a puff of air to the cornea of the left eye, was delivered by the fall of an 80-mm. column of mercury contained in a U tube. The moment of onset of the air puff, 1.50 sec. after onset of the CS, was photographically recorded.

In an attempt to produce consistent interstimulus intervals with slow swings of the pendulum, it was necessary for the pendulum and the photographic film rack (mounted on the pendulum arm) to pass the lighted aperture at a rate too rapid to allow recording of the whole interstimulus interval. Thus the first .50 sec. following the onset of the CS were not recorded. On the basis of findings generally obtained with trace conditioning procedures where *S*s appear to learn to delay CR's until just before the presentation of the UCS, it was anticipated that recording only the last 1 min. of the interstimulus interval would adequately show the relative levels of conditioning obtained. The data appear to confirm this expectation.

Silent motion pictures were shown *S*s during the experimental sessions to distract them from the purpose of the study, and to keep them awake during the long experimental sessions. An Ansco 16-mm. projector was used to present educational films spliced together and wound on 1600-ft. reels. A 400-ft. reel was also presented to each member of the distributed group during the preliminary test trials.

The films were projected on a rectangular milk-glass screen, 8 by 10 in., 28 in. in front of *S* when in the "ready" position and about 40 in. in front at other times. The base of the screen was 3 in. above the eye level of *S*s. The projector was placed behind the screen, facing *S*, and adjusted so that the movie covered the whole screen.

Subjects.—Twenty-two *S*s, 14 male and 8 female, were employed in the experiment. Each of the experimental groups consisted of seven male and three female *S*s, and two female *S*s were used as controls. (It was not possible to utilize the records of two male experimental *S*s, one in each group, who blinked continuously on a large proportion of recorded trials.) All were students in the introductory psychology course at the State University of Iowa.

Procedure.—The *S*s were told that the purpose of the experiment was to measure eye fatigue involved in watching motion pictures, which would be indicated by the latency and amplitude of blinks to the air puff. They were instructed to ignore an occasional tone heard through the earphones which were employed to help muffle the noise from the projector. When *E* said "ready," *S*s were to lean forward and bite the biting board until a photograph was taken; they were then to sit back.

Both experimental groups were given three test trials to the CS alone preliminary to the joint presentation of the stimuli. Approximately 1 min. was allowed to elapse after the film was started before the first test stimulus was presented. The intertrial interval for these test trials was a constant 4 min. for both groups. The mean intertrial interval for Group I during conditioning was 270 sec. with a range from 180 to 360 sec.; for Group II, 120 sec. with a range from 80 to 160 sec.

The *S*s in these groups were trained through four daily sessions separated by 20 to 28 hr. The three test trials and 20 reinforced trials were given on the first day; 20 more reinforced trials were given on each of the three succeeding days. Thus, there were 80 acquisition trials in all. The daily sessions for Group I lasted approximately 2 hr.; for Group II, approximately 1 hr.

Since the reel was changed for Group I after the test trials, and between Trials 19 and 20 or Trials 20 and 21 of each daily session, the projector was also shut off for Group II for 1 min. between these same trials. The projector was allowed to run for at least 45 sec. before the next trial. In no case did the reel change plus the 45-sec. adaptation require more time than the prescribed intertrial interval. All reels were presented equally often on each day of training.

Owing to lack of time, it was possible to run only two control *S*s. They were utilized to obtain an indication of the rate of blinking due to factors other than the joint presentation of the stimuli. The conditions for these two *S*s were identical with those of the experimental *S*s except that the tone was never presented. For one control *S*, the intertrial interval was 270 sec., for the other, 120 sec. The instructions to the controls differed from those to the experimental groups only in that no mention was made of the tone.

Records were taken on the three test trials, and on acquisition Trials 1, 5, 9, 10, 19, 20, 21, 25, 29, 30, 39, 40, and 41 through 80. In order to

FIG. 1. Acquisition curves for Groups I and II when the criterion included every response at least .5 sec. in latency and 1 mm. in amplitude

avoid differential cues from trial to trial, a sheet of exposed film was placed in the holder on the trials on which no records were taken.

Response measure criteria.—The response measure utilized is the per cent of records on which there occurred at least one eyelid closure. It will be recalled that responses during the first .50 sec. following the onset of the CS were not recorded. Thus, for three different criteria employed in counting the blinks, only those with a latency of .50 sec. or longer were included. The first criterion included all responses with an amplitude of at least 1 mm.; the second included responses of 2 mm. and larger; the third required response amplitudes of 5 mm. Since the curves obtained with these criteria are essentially the same, only the first set of data is presented.

RESULTS AND DISCUSSION

The eyelid responses obtained in the present study were probably of two kinds: (*a*) those of an unlearned variety, including random blinks and blinks resulting from the presentation of the motion pictures, and (*b*) conditioned responses. As a result, the curves presented contain all the responses meeting the criteria used in reading the records.

Figure 1 represents the per cent of responses obtained per block of trials.[5] That learning actually did occur under the experimental conditions is indicated, first of all, by the fact that the mean per cent of responses for the control *S*s on all recorded trials was 48.2 as compared with a mean per cent of 73.1 for Group I, and of 66.5 for Group II. In addition, the curves, although irregular (perhaps owing to the exclusion of responses less than .50 sec. in latency), show a general rise until the 41–50 trial block where both groups appear to have reached a maximum. There was no comparable increase in the case of the controls. The mean per cent of response for the control *S*s for the first 15 records obtained (through Trial 40) was 53.3 as compared with a mean per cent of 46.2 for the last 40 trials.

Before the conditioning procedure began, Group I responded more frequently than Group II. This is shown in Fig. 1 by a comparison of the initial points for the two groups. In addition, Group I also reached a somewhat higher maximum response frequency than Group II in the last four blocks of trials. The superiority of Group I is not at all significant in either case. A test of significance of the difference in response frequency prior to paired presentation of the stimuli yielded a t of 1.08 ($df = 16$). When the responses obtained for the two groups on the last 40 trials were compared, a t of .682 was obtained. Considering the large variability within each of the groups, it is estimated that large samples would be necessary to obtain statistical significance if the

[5] The first point plotted represents the per cent of records on which a response meeting this criterion was found on the first three test trials and the first acquisition trial; the second point, the per cent of records on Trials 5, 9, and 10; the third, on Trials 19, 20, and 21; the fourth, Trials 25, 29, and 30; the fifth, Trials 39 and 40, and the remaining four points, on successive blocks of ten trials each.

real difference is as small as the one obtained.

It is apparent, then, that conditioning occurred in Groups I and II, and that the difference between the per cent of *total* responses on the last 40 acquisition trials for the two groups was not significant. The question still remains as to whether the maximum frequencies of *conditioned* responses are different. It was indicated above that there was little change in the frequency of unlearned responses for the control *Ss*; it would appear reasonable to infer that there was also no change in unlearned responses for either experimental group throughout training. Since the nonsignificant initial difference in frequency of *unlearned* responses probably remained nonsignificant for both groups of experimental *Ss* at the end of training, it follows that the lack of significance of the difference between the maximum per cent of *total* responses may be interpreted as a nonsignificant difference between the maximum per cent of *conditioned* responses made by the 270- and 120-sec. intertrial interval groups. Finally, since no significant difference in conditioning was found between the two long intertrial interval groups, it may be inferred that the function relating the intertrial interval and the limit of conditioning approaches an asymptote. In view of the results reported by Spence and Norris (8), the data are interpreted to indicate that this function is probably negatively accelerated in shape.

Summary

Trace conditioning procedures to a tone were employed with 18 *Ss*, nine with an intertrial interval of 270 sec.; for the other nine, 120 sec. The interval between the CS and UCS (a puff of air to the left cornea) was 1.50 sec. Both groups were given 80 reinforced trials equally distributed over four experimental sessions, during which motion pictures were presented. Two control *Ss* were utilized; the conditions for these *Ss* were the same as those for the experimental groups except that the conditioned stimulus was not presented. The intertrial interval for one of these *Ss* was 120 sec.; for the other 270 sec.

The results showed no significant difference in the maximum amount of conditioning obtained with these two intertrial intervals.

It was concluded that the function relating length of the intertrial interval and the limit of conditioning reaches an asymptote, and in view of other results (8) is probably negatively accelerated.

(Received June 9, 1952)

References

1. Dodge, R. A pendulum-photochronograph. *J. exp. Psychol.*, 1926, **9**, 155–161.
2. Hilgard, E. R. Modification of reflexes and conditioned reactions. *J. gen. Psychol.*, 1933, **9**, 210–215.
3. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: D. Appleton-Century, 1940.
4. Humphreys, L. G. Distributed practice in the development of the conditioned eyelid reaction. *J. gen. Psychol.*, 1940, **22**, 379–385.
5. Kimble, G. A. Conditioning as a function of the time between conditioned and unconditioned stimuli. *J. exp. Psychol.*, 1947, **37**, 1–15.
6. Reynolds, B. The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. *J. exp. Psychol.*, 1945, **35**, 15–30.
7. Schlosberg, H. Conditioned responses in the white rat. *J. gen. Psychol.*, 1934, **45**, 303–335.
8. Spence, K. W., & Norris, E. B. Eyelid conditioning as a function of the inter-trial interval. *J. exp. Psychol.*, 1950, **40**, 716–720.

INDIVIDUAL DIFFERENCES IN DEFENSIVE FORGETTING

CHARLES W. ERIKSEN

The Johns Hopkins University

Studies of memory for completed and incompleted tasks have generally shown that under ego involvement *Ss*, as a group, tend to favor completed or successful tasks in their recall as opposed to incompleted or failed tasks (**5, 6, 8**). This finding has frequently been interpreted as demonstrating defensive forgetting. It has been assumed that the threat to the ego or self-esteem posed by failure in ego-involved areas is sufficient to bring a defensive mechanism into play.

With but few exceptions (**1, 5**), it has been implicitly assumed that all people respond to psychological threat in a similar way, and *Ss* who show little or no tendency to favor completed tasks in their recall have generally been overlooked. If the completed task method is to be considered a valid measure of defensive forgetting, then these deviant individuals must be considered the result of either the crudeness of the method or of individual differences in response to ego threat. Assuming the latter explanation to be correct, we should then expect these individual differences in response to appear in other situations where ego threat is involved. In addition, if the incompleted task method is to have any generality, we should expect individuals to show consistency in their response to different threatening situations.

Clinical as well as experimental evidence suggests that people are consistently different in their response to threat. In a previous experiment (**4**) it was reported that *Ss* who showed a marked tendency to favor completed tasks in their recall also had relatively high recognition thresholds (tachistoscopically determined) for words with long association times. On the other hand, *Ss* who favored the failed tasks in their recall had little or no difficulty in recognizing such words. The present study is a further investigation of individual differences in response to ego threat. Differential learning rates for words with long, medium, and short association times have been investigated by the paired-associates method in two groups of *Ss*: (*a*) those who tend to recall completed tasks under ego involvement, and (*b*) those who tend to recall failed tasks.

PROCEDURE

Experiment I.—This was a study of memory for completed and incompleted tasks. The tasks consisted of 16 20-word sentences which had been scrambled into two- and three-word phrases. The *S* was required to organize the phrases into a meaningful sentence. Eight of the sentences were solvable and the rest unsolvable. To facilitate recall each sentence was given a title by underlining one of the two-word phrases. The sentences were administered in group form by means of a slide projector. As each sentence appeared on the screen, *E* read aloud the title. Sentences were shown for 3 min. each during which time *S* was to write his solution on separate pages of an examination booklet. The order of appearance of solvable and unsolvable sentences was arranged so as to minimize serial-position effects.

The sentences were administered to the experimental (ego-involved) *Ss* under the guise of an intelligence test. They were told that *E* was interested in selecting *Ss* who were bright enough to participate in an experiment on abstract reasoning. They were further told that anyone with an IQ of 115 or better should have little or no difficulty in solving all of the sentences within the time limits. To increase the deception, classmate stooges were scattered throughout the group. These stooges pretended to complete successfully all of the 16 sentences.

The sentences were administered to control *Ss* under identical conditions except for different

instructions and the absence of stooges. They were told that the sentences had been made up by *E* to keep *S*s in a learning experiment occupied during rest intervals so they could not rehearse, and that *E* simply wanted to determine if the sentences were difficult enough to keep *S*s busy for the necessary lengths of time. The control group was further told that *E* was already certain that many of the sentences were too difficult to be solved within the time interval. The *S*s were asked to work as rapidly as possible in order to make their performance comparable to that of the learning *S*s.

The scrambled sentences were followed by a 5-min test of "word fluency" which consisted of writing down four-letter words that began with the letter "t." A 5-min. recall period followed in which *S*s wrote down the titles or the content of the sentences they could remember.

All *S*s were male undergraduate students enrolled in an introductory course in psychology. Seventy-six *S*s comprised the experimental group and 36 the control. However, 13 *S*s in the control and 32 *S*s in the experimental group were discarded because of failure to complete all of the solvable sentences.

Experiment II.—The *S*s for this experiment were selected from among the experimental group in Exp. I. For each *S* in the experimental group a ratio was computed by subtracting the number of completed tasks recalled from the number of incompleted and dividing by the total tasks recalled (I−C/I+C). On the basis of these ratios the ten *S*s who showed the strongest tendency to favor incompleted tasks in their recall (Failure Recall Group) and the ten who showed the strongest tendency to favor completed tasks (Success Recall Group) were chosen for Exp. II. Each of these 20 *S*s was administered, individually, a 95-item word-association list. This list was identical with the list reported by Bruner and Postman (2) except for the following words which were deleted: hymen, pubis, and *S*'s first and last name. Eight words with the longest, eight with the shortest, and eight with medium association times were chosen for each *S* on the basis of his distribution of association times. A comparison of the word-association test performance of the Success and Failure Recall Groups showed no significant difference in either average association time or in variability of association time.

Paired associates were formed by using nonsense syllables as stimuli and words as responses. Twenty-four three-letter nonsense syllables of low association value were divided into three groups of eight each. Group 1 syllables were always paired with an *S*'s long association-time words, Group 2 with his medium association-time words, and Group 3 with his short association-time words. The nonsense syllables were typed individually on the face of 3 by 5-in. cards and on the reverse side of the card the nonsense syllable appeared with its paired word. Each *S* was presented with his pack of 24 cards. At a signal from *E*, *S* removed the blank top card and spelled aloud the nonsense syllable of the first paired associate. At the end of 5 sec., he turned the card over and spoke the response word. Another 5 sec. was allowed for studying the response word. Pacing was by means of signals coming from a tape recorder. In order to control overlearning, pairs were dropped from the list when the response word had been correctly anticipated three out of four times. Learning was continued until all pairs had reached this criterion. Twenty-four hours later a relearning session was held and the list was again learned to the criterion. The learning sessions were individually administered. During learning and relearning the order of the pairs within the list was randomized between trials.

Results

Table 1 compares the experimental and control groups of Exp. I on the recall of completed and incompleted tasks. Three different comparisons are provided in this table. When the two groups are compared on the difference score obtained by subtracting for each *S* the number of completed tasks recalled from the number of incompleted recalled, it is found that the experimental *S*s show a significant difference in the direction of more completed tasks recalled. When the

TABLE 1

Number of Completed and Incompleted Tasks Recalled by the Ego-Involved and Control *S*s

Tasks	Ego-Involved ($n = 44$)		Non-Ego-Involved ($n = 23$)		t*
	Mean	*SD*	Mean	*SD*	
Incomplete	2.0	1.2	3.4	1.4	4.29
Complete	3.6	1.5	3.7	1.1	.28
Difference (I−C)	−1.6	2.0	− .3	1.9	2.52

* A t of 2.39 is significant at the .01 level for a one-tailed test.

TABLE 2

DIFFERENCES IN LEARNING AND RELEARNING RATE FOR LONG, MEDIUM, AND SHORT ASSOCIATION-TIME WORDS FOR THE SUCCESS AND FAILURE RECALL GROUPS

Association Times of Words	Failure Recall Group						Success Recall Group						t* Group Difference	
	Learning			Relearning			Learning			Relearning				
	Mean Diff.	*SD* Diff.	t	Mean Diff.	*SD* Diff.	t	Mean Diff.	*SD* Diff.	t	Mean Diff.	*SD* Diff.	t	L	RL
Long-Short	.24	.79	.96	−.01	.97	.00	2.47	2.37	3.29	2.26	2.19	3.26	2.82	3.01
Long-Medium	−.05	1.63	.10	.13	.93	.44	2.30	1.48	4.91	1.51	1.71	2.79	3.37	2.24
Medium-Short	−.15	1.51	.31	−.29	.95	.96	.17	1.59	.34	.75	1.49	1.59	.46	1.86

* t has been corrected for heterogeneity of variance via the method suggested by Cochran and Cox (3). t values of 1.83 and 2.82 are significant at the .05 and the .01 levels for a one-tailed test.

groups are compared on the number of completed and the number of incompleted tasks recalled, it is found that both groups recall the same number of completed tasks but the experimental group recalls significantly fewer incompleted ones.

For the paired-associates learning of Exp. II the mean number of trials to learn and relearn the long, medium, and short association-time words was obtained for each *S*. A modified four-way analysis of variance was performed on these data based upon the classification into Success and Failure Recall Groups, association time, learning and relearning, and individuals. This analysis showed no over-all difference in learning between the two groups but did reveal a highly significant interaction ($p = .001$) between Success and Failure Recall Groups and association times.

In order to evaluate this interaction, three learning and three relearning difference scores were obtained for each *S* in the two groups. For each *S* the mean number of trials to learn the long, medium, and short association-time words was computed. Then difference scores were obtained by subtracting the mean of his learning trials for medium and for short association-time words from the mean for the long association-time words and also his mean trials for the short from the medium. The same procedure was used on the relearning trials.

The mean difference scores obtained for the learning and relearning trials are presented in Table 2. For the Failure Recall Group none of the mean differences is significantly different from zero. For this group there is no difference in rate of learning or relearning for words with long, medium, or short association times. For the Success Recall Group, on the other hand, long association-time words require significantly more trials to learn and to relearn than do either the medium or short association-time words. However, there is no significant difference in learning rate between medium and short association words. Evaluation of the differences between the mean difference scores for the Success and Failure Recall Groups substantiate these within-group comparisons.

DISCUSSION

The above results have two important implications. First, these results support the view that defensive forgetting does occur in the completed-incompleted tasks experiment under conditions of ego-involvement. The results of previous investigators (5, 6,

8) are confirmed by the finding that, individually, ego-involved *Ss* tend to recall more completed than incompleted tasks as compared with non-ego-involved *Ss*, and that as a group they recall significantly less incompleted tasks than does the non-ego-involved group. More important to an hypothesis of defensive forgetting, however, is the finding that performance in the completed-incompleted task experiment is related to other situations where defensive forgetting might be expected to occur. The *Ss* who show evidence of defensive forgetting in the completed-incompleted task situation are found to require more trials to learn and relearn affective words than neutral words. On the other hand, *Ss* who show little or no evidence of defensive forgetting in the completed-incompleted task experiment learn affective words at least as easily as they do the neutral. This consistency in response in two independent situations provides strong evidence for the defensive forgetting hypothesis.

The second implication of the above data concerns the difference in responses to ego threat shown by the Success and Failure Recall Groups. Whereas the Success Recall Group gives evidence of defensive forgetting in both experiments, the Failure Recall Group shows little or no indication of defensive forgetting in either experiment. Whether this difference in response to ego threat, as shown by the Success and Failure Recall Groups, is of a quantitative or qualitative nature, cannot be determined from the present data. It does suggest, however, that any experiment involving psychological defenses must be prepared to take account of this difference. Perhaps the indifferent success achieved by the numerous experiments on rate of learning and recall of pleasant and unpleasant, neutral and affective material (7) can be explained on the basis of such individual differences in response to ego threat.

Summary

The present study was concerned with the relation between memory for completed and incompleted tasks and the learning of neutral and affective words. Two subexperiments were involved. Experiment I was a study of memory for completed and incompleted tasks. Sixteen scrambled sentences (eight solvable and eight unsolvable) were administered to an experimental (ego-involved) group and a control (non-ego-involved) group. Experiment II was a study of paired-associates learning in which three-letter nonsense syllables were paired with words having long, medium, and short association times for individual *Ss*. The *Ss* in this experiment were selected from the ego-involved group of Exp. I on the basis of their having recalled relatively more completed or relatively more incompleted tasks.

The results show that in Exp. I, while both groups recall the same number of completed tasks, the experimental group recalled significantly fewer incompleted tasks. In Exp. II *Ss* who favored completed tasks in their recall required significantly more trials to learn and to relearn words with long association times than to learn or relearn words with either medium or short association times. The *Ss* who favored incompleted tasks in their recall, on the other hand, learned words with long association times at least as readily as those with medium or short association times. Differences between the high and low groups were significant. The results support an hypothesis of defensive forgetting and in addition show

marked and consistent individual differences in response to ego threat.

(Received June 16, 1952)

References

1. Alper, T. G. Memory of completed and incompleted tasks as a function of personality: An analysis of group data. *J. abnorm. soc. Psychol.*, 1946, **41**, 403–421.
2. Bruner, J., & Postman, L. Emotional selectivity in perception and reaction. *J. Pers.*, 1947, **16**, 69–77.
3. Cochran, W. G., & Cox, G. M. *Experimental designs*. New York: Wiley, 1950.
4. Eriksen, C. W. Defense against ego threat in memory and perception. *J. abnorm. soc. Psychol.*, 1952, **47**, 230–236.
5. Glixman, A. F. Recall of completed and incompleted activities under varying degrees of stress. *J. exp. Psychol.*, 1949, **39**, 281–295.
6. Lewis, H. B., & Franklin, M. An experimental study of the role of the ego in work. II. The significance of task orientation in work. *J. exp. Psychol.*, 1944, **34**, 195–212.
7. Rapaport, D. *Emotions and memory*. New York: International Universities Press, 1950.
8. Rosenzweig, S. An experimental study of repression with special reference to need-persistive and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, **32**, 64–74.
9. Rosenzweig, S., & Sarason, S. An experimental study of the triadic hypothesis: Reaction to frustration, ego defense, and hypnotizability. I. Correlational approach. *Charact. & Pers.*, 1942, **11**, 1–20.

TEMPORAL NUMEROSITY: I. PERCEIVED NUMBER AS A FUNCTION OF FLASH NUMBER AND RATE[1]

PAUL G. CHEATHAM AND C. T. WHITE

U. S. Navy Electronics Laboratory

Taubman, as reported in a recent paper (6), investigated the judgment of visual number as a function of the number of flashes of light and the time interval between flashes. For intervals ranging from 500 msec. to 143 msec., and numbers from 1 to 10, there was a consistent tendency toward underestimation of the stimulus number. This underestimation increased for higher numbers and shorter intervals. Similar results have been obtained for the judgment of auditory number (3, 5).

The highest flash rate used by Taubman was approximately 5/sec., a rate at which *O* might conceivably be able to count, with some degree of accuracy, all the flashes presented. It would be of interest to know how *O* would react to higher rates and larger numbers in which case one would not expect him to be able to count accurately. An extension of Taubman's curves would indicate that estimation of stimulus number would decrease in a linear fashion. Of special interest, moreover, would be the degree of estimation at flash rates of about 10/sec., the region at which brightness enhancement occurs (2).

The general purpose of this experiment was to investigate perceived number as a function of flash rates higher than 5/sec.

METHOD

The method consisted of presenting to *O* groups of flashes, varying in number, at different repetition rates. The flash series was produced by a sectored disk driven by a constant-speed synchronous motor, geared to 1 rps. A separate disk was used for each repetition rate. The sectors were cut to produce a constant 11-msec. flash duration, with dark intervals of 22, 33, 55, and 88 msec. corresponding to rates of 30/sec., 22.5/sec., 15/sec., and 10/sec. The number of flashes was varied by exposing the proper number of sectors around the disk. An auxiliary shutter provided control of intertrial intervals.

A large (17-cm. diameter) uniformly illuminated surround of 1.8 ft.-L. served as an adapting field, at the center of which the flash appeared in a circular test field (.6 cm. diameter). The observation distance was 80 cm. The test field was illuminated by the beam of a 100-w. projector system, which was brought to a sharp focus in the plane of the sectored disk. The brightness of the test field, when illuminated by the uninterrupted light beam, was 1400 ft.-L. This illumination was held constant throughout the experiment.

The *O* was seated in front of the stimulus field, shielded from *E* by a large black partition, and allowed to adapt for 10 min. to the level of the surround. He was instructed to fixate the test field at the ready signal, 2 sec. before the flash series was presented. No fixation point was required since the circular contour of the test field was clearly visible. The *O* was required to report the number of flashes perceived.

The flashes were presented by the method of constant stimuli. For any given rate, the number of flashes presented was randomized. Rates, in turn, were changed randomly. Each stimulus number was presented ten times at all rates to each of the five *O*s. Experimental sessions lasted approximately 1 hr.

RESULTS

The data are given in Table 1 and are plotted in Fig. 1. Two major facts are obvious. (*a*) Underestimation of the stimulus number was marked in all cases. The *O*s never perceived the objective number correctly, except in the case of two or three flashes at the lower rates. For this reason, analysis in terms of degree of accuracy would be of little

[1] U. S. Navy Electronics Laboratory Professional Contribution No. 2.

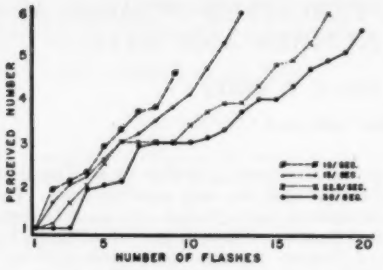


FIG. 1. Number of flashes perceived as a function of the number of flashes presented. Each point represents the mean of 50 responses (10 responses for each of 5 *Os*).

value. (*b*) Estimation of stimulus number did not decrease in a linear fashion. The empirical curves in Fig. 1 show definite plateaus. Furthermore, responses to certain stimulus numbers are extremely stable, having little or no variation around the mean value (Table 1). Because of the apparent stepwise characteristics of the curves and the stability of certain responses, no attempt was made to fit straight lines to the data.

Of special interest are the stable points. These would indicate that certain numbers of objective flashes were grouped into very stable perceptual units. For example, 8, 9, or 10 flashes at 30/sec. were perceived by all observers on all trials as three flashes.

Portions of curves between stable points represent some type of transition from one perceptual unit to the next. These appear to be of a probability-type function in which the addition of one objective flash increases the probability that *O* will perceive one more subjective flash.

The same data are presented in a different way in Fig. 2. In this case the abscissa represents the cumulative time from the beginning of a train of flashes, and each point indicates a flash. (In effect this graph presents perceived number of flashes as a func-

TABLE 1

DISTRIBUTION OF RESPONSES TO VARYING NUMBERS OF FLASHES AT VARIOUS RATES OF PRESENTATION

Number of Flashes	30/sec. Rate						22.5/sec. Rate						15/sec. Rate						10/sec. Rate					
	Perceived Number						Perceived Number						Perceived Number						Perceived Number					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	50						50						50						50					
2	50						45	5					20	30					6	44				
3	49	1					20	30						50						45	5			
4	7	43						50						38	12					33	17			
5	1	49						23	27					19	31					6	44			
6		43	7					2	48					1	49						37	13		
7		6	44						50						40	10					17	33		
8			50						50						27	23					6	44		
9			50						48	2					9	41						19	31	
10			50						30	20						47	3							
11			46	4					14	36						13	37							
12			37	13					4	46							35	15						
13			16	34					3	47							2	48						
14			1	49						37	13													
15				50						13	36	1												
16				37	13					3	47													
17				17	33						36	14												
18				5	45	1						50												
19					47	3																		
20					22	28																		

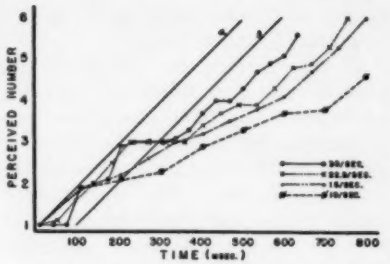


FIG. 2. Number of flashes perceived as a function of the exposure duration. Each point represents a stimulus flash. (Lines *a* and *b* are explained in the text.)

tion of the duration of exposure of a light source flashing at the various rates.) This method of presenting the data is enlightening in that it shows that perceived number is not only a function of stimulus number but is also dependent on the total elapsed time. Thus a train of eight flashes at 30/sec. and one of six flashes at 22.5/sec. require the same amount of time and are both perceived as three flashes, whereas six flashes at 30/sec. are seen as two.

DISCUSSION

The results obtained in this experiment indicate that the technique employed is a valuable tool in the study of the basic temporal aspects of vision.

The most surprising result was the subjective grouping of stimulus flashes into perceptual units. This grouping was consistent for all *O*s. (The subjective flashes were reported as large pulses of light, expanding and contracting with a gamma-type apparent motion.) Counting these pulses was quite simple. The *O*s reported their judgments to be based on actual counting of perceptual flashes and not on any sort of estimation of physical flashes.

The highest perceived number reached in this study was six flashes. Extrapolation of the curves in Fig. 2, up to 1-sec. cumulative time, suggests that the subjective rate did not exceed 6–8/sec. although the objective rate was varied from 10–30 flashes/sec. This would seem to indicate that the subjective rate is determined primarily by some temporal patterning in the visual process and not by the objective flash rate.

The reliability and the temporal pattern of the responses are the most important aspects of the results from a theoretical point of view. It is apparent that we are not dealing with a judgment process, but rather with some basic property of vision itself. On the basis of these results, and on pilot studies done at rates as low as 5/sec., the conclusion is inescapable that even at these low rates certain flashes are "lost"; therefore the low readings Taubman obtained perhaps should not be regarded as "underestimation" in the strict sense of the word.

One basic question to be asked at this point is: What flashes are "lost," and by what mechanisms do they become "lost"? On the basis of a short pilot study, as well as by inspection of the data presented, one may hypothesize a definite interaction effect between flashes. For example, four flashes at 30/sec., and two flashes at 10/sec., are seen as two. The exposure interval or total time in both cases is 110 msec. But, the 88-msec. interval between the first and second flash at 10/sec. is objectively dark, whereas this same interval between the first and fourth flash at 30/sec. is filled with two light-dark cycles. Apparently, then, the two flashes in between are not needed for the numerosity or *counting* process. Their presence would, on the other hand, produce changes in the subjective

brightness discrimination. The mechanism operating here must be either a facilitative or inhibitory one. A study using brightness discrimination would indicate summation, if the four flashes at 30/sec. appeared brighter than the two flashes at 10/sec., or inhibition, if the four appeared less bright. Selective adding and removing of different intermediate flashes would show which flashes are being most affected by the interaction process operating.

The results obtained by the present method are even more interesting in that they can be a perceptual reference for comparable electrophysiological studies. It is known, for example, that an intermittent light can produce "ripples" on the electroretinogram (**1, 4**). The correlation of these "ripples," or some other aspect of the human ERG, with perceived number might serve the dual purpose of adding to our knowledge of the visual process and enhancing the usefulness of the ERG in visual research. Such a study is now under way. Another approach to the problem could be by means of the EEG, especially for studying the possible relationship between the temporal pattern which was found and the alpha rhythm or some other salient feature.

A mathematical analysis of the data will not be presented in the present paper. However there is one general observation to be made that is quite suggestive. In Fig. 2, line *a* represents "perfect" response to 10 flashes/sec. It can be seen that the slope of this line is the same as the over-all slope of the points for the higher flash rates —at least for the first 300 msec. of exposure duration. Following the plateau which occurs in this region, it can be seen that the slope of the highest flash rate again approaches that of the "perfect" 10/sec. rate. Line *b* has been put in to help show this point. It has the same slope as *a*, but is displaced 100 msec. in time.

On the basis of the results of this experiment, the following working hypothesis can be stated: The perceived number of a series of flashes is limited by the total time taken by the series of flashes. The upper limit seems to be about one perceived flash per 100 msec. The more objective flashes in each 100-msec. unit of time, the nearer the upper limit is approached—as long as the flash rate remains below fusion frequency.

The above holds, of course, only for the first 300 msec. If it held throughout the range, all the points would fall between *a* and *b*. A complete hypothesis should account for the displacement at 300 msec. It is hoped that experimentation now under way may help to develop such an hypothesis.

Summary

The perceived number of flashes was studied as a function of the flash rate and the number of flashes presented. Throughout the experiment each flash was of constant duration and brightness (11 msec.; 1400 ft.-L.), differences in rate being accomplished by changing the interval between flashes. Dark intervals of 22, 33, 55, and 88 msec. were used, corresponding to rates of 30, 22.5, 15, and 10 flashes/sec., respectively.

The results show that the subjective flash rate probably will not exceed 6–8/sec., regardless of the objective flash rate. Each successive subjective flash is not added in an equal unit of time, but a definite temporal pattern appears. On the basis of the present results the conclusion is made that the perceived number is dependent on both the number of flashes presented and the total time in which they are presented. The total time imposes an upper limit to perceived number,

and the number of flashes presented determines how closely this limit is approached. The upper limit appears to be about one subjective flash per 100-msec. elapsed time, but this function is discontinuous, a break occurring in the region of 300 msec.

(Received June 30, 1952)

References

1. Bartley, S. H. Some observations on the organization of the retinal response. *Amer. J. Physiol.*, 1937, **120**, 184–189.
2. Bartley, S. H. *Vision.* New York: Van Nostrand, 1941.
3. Garner, W. R. The accuracy of counting repeated short tones. *J. exp. Psychol.*, 1951, **41**, 310–316.
4. Granit, R., & Riddell, L. A. The electrical responses of light- and dark-adapted frogs' eyes to rhythmic and continuous stimuli. *J. Physiol.*, 1934, **81**, 1–29.
5. Taubman, R. E. Studies in judged number: I. The judgment of auditory number. *J. gen. Psychol.*, 1950, **43**, 167–194.
6. Taubman, R. E. Studies in judged number: II. The judgment of visual number. *J. gen. Psychol.*, 1950, **43**, 195–219.

FIGURE–GROUND CONTRAST AND BINOCULAR RIVALRY[1]

L. T. ALEXANDER AND P. D. BRICKER

The Johns Hopkins University

When the two halves of a stereogram contain figures which are different, the figures may fuse and produce the impression of depth. If the figures are made sufficiently different, they do not fuse. Instead, they alternate so that first one and then the other is seen. In a previous experiment (1) it was shown that the rate at which such figures alternate is higher if (*a*) the figure contour is continuous than if it is broken, and (*b*) there is greater brightness contrast between the figure and its ground. This study is an attempt to describe more fully the rate-contrast relationship.

METHOD

Apparatus.—Binocular rivalry was produced by presenting to *S* split stereograms in a Renshaw Stereo-Disparator. This instrument and the method for recording and measuring rivalry rate have been described elsewhere (1) and will not be repeated. Five stereograms were used. The left half of each consisted of a series of black vertical bars and the right half of a similar series of black horizontal bars. The bars were drawn in India ink on photographic paper which had been exposed and processed beforehand so that each stereogram had a background of a different shade of grey. These greys were of 10%, 30%, 50%, 70%, and 90% reflectance as measured by the Macbeth Illuminometer. Since the black bars had a reflectance of 1.4%, the five stereograms represented figure-ground contrast ratios of .86, .95, .97, .98, and .98, respectively. In the remainder of this report each stereogram will be referred to by the reflectance of its background, i.e., "10," "30," etc.

Subjects.—Five male undergraduate students at The Johns Hopkins University served as *Ss*. Since there is some evidence that binocular rivalry rate may be related to certain other visual skills, *Ss* were selected to meet the following conditions: less than 4 diopters lateral phoria at far point; 100% far point acuity without correction; a minimum stereo-acuity of 19 sec. The first measure was obtained using the Stereo-Disparator, and the other two were obtained on the Bausch and Lomb Ortho-Rater.

Procedure.—The following instructions were read to *S:* "When you look into the eyepieces, you will see a series of vertical lines and a series of horizontal lines. In the center of the field there will be a cross. You will have three or four minutes to become acquainted with the experimental situation and to familiarize yourself with the phenomenon of binocular rivalry. After a while you will notice that the vertical and horizontal lines are not seen superimposed as a lattice, but that they appear and disappear alternately. Fixate the cross in the center of the field and report by pressing the key each time the cross appears between vertical lines. When the cross no longer appears between vertical lines, release the key. It may be that the whole field will not alternate as a unit. You should be concerned only with the region around the cross."

Each *S* was given a 1-min. practice trial with each stereogram. Forty-eight hours later he was given 20 1-min. practice trials with the "90" stereogram. This was done to show *S* the stereograms to be used and then to give him experience in observing binocular rivalry. Forty-eight hours later the main experiment was begun.

The stereograms were presented in accordance with a multiple latin-square design with different latin squares drawn at random (2) for each *S*. Each *S* observed each stereogram five times, a total of 25 observations. A trial was 1 min. and 10 sec. long, the first 10 sec. of which were not included in the data. After each trial there was a 3-min. rest, except after Trials 10 and 20, at which time there were 10-min. rests.

RESULTS

Individual analyses.—A separate analysis of variance was performed for each *S*. With these it was possible to account for the variance contributed by the contrasts, the sequences of contrasts, and the contrast order within

[1] This research was carried out under Contract N5-ori-166, Task Order I, between the Systems Coordination Division, Naval Research Laboratory, Office of Naval Research, and The Johns Hopkins University. This is report No. 166-I-158, Project Designation No. NR 507–470, under that contract.

sequences. Contrasts reached at least the .05 level of confidence in four of the five *S*s. Orders accounted for an insignificant amount of the variance, but the variance contributed by sequences approached significance in two *S*s.

In spite of the latter outcome it was decided to combine the data for all *S*s and to treat all trials on the same contrast for each *S* as replications. With this procedure sequence variance is confounded with replication variance and the error term is inflated. Although this procedure will not give an accurate picture of the relationship between contrast and rivalry rate, the reliability of any differences will be underestimated. That is, the variance contributed by the different contrasts must be greater in order that an acceptable confidence level be reached.

Combined analysis.—Table 1 shows the results of the analysis in which the data for all *S*s are combined. It can be seen from this table that the contrast variance is clearly significant. In spite of efforts to restrict inter-*S* variability by selecting *S*s in accordance with scores on other presumably related visual tasks, *S* variance is also highly significant.

Figure 1 shows the average rate of alternation with each stereogram. The abscissa is a scale of contrast ratios on which the five stereograms are identified by their background reflectances. The letters refer to individual *S*s.

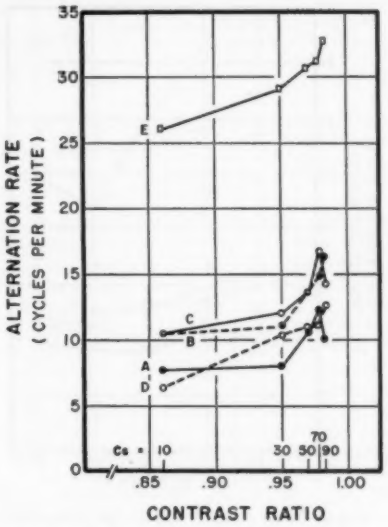


FIG. 1. Rivalry rate as a function of figure-ground contrast for each *S*. Each point is the mean of five observations.

TABLE 1

DOUBLE-CLASSIFICATION ANALYSIS OF VARIANCE OF RIVALRY RATE SCORES

Source of Variation	Sums of Squares	*df*	*F*
Contrasts	469.1	4	7.68*
Subjects	6958.4	4	113.91*
Subjects × contrasts	79.7	16	
Error	1527.2	100	
Total	9034.4	124	

* Significant beyond the .001 level of confidence.

In order to test the hypothesis that the mean rate differences obtained were due to sampling errors, *t* tests were run for all combinations of points for each *S*. Of the 50 possible tests 30 were below the .10 level of confidence, six were between .10 and .05, seven were between .05 and .01, six were between .01 and .001, and one was less than .001. Three of these were reversals. Although there is the suggestion of a systematic relationship here, the reliability of such a relationship is quite low.

In a previous experiment on rivalry (1) there was very little intertrial variability for any of the *S*s on any contrast. In all cases the mean difference between contrasts was greater than four times its standard error. The present data show much higher inter-

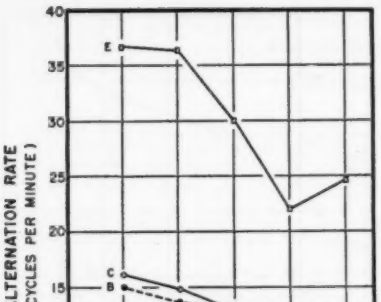


FIG. 2. Mean rivalry rate for successive blocks of five trials for each *S*. Each point includes all five contrasts.

trial variability. It is possible that massing of trials contributed to the present results since in the previous experiment all trials were at least 4 hr. apart.

There is an indication that some sort of "fatigue" effect is operating to decrease the rate of alternation as the number of observations increases. Figure 2 shows this decrease in rivalry rate as a function of successive blocks of five trials. All five stereograms are included in every block. The data for each *S* are plotted separately. It appears that the higher the initial average rate, the more rapid is the decrease in rate.

It seems obvious that in order to be able to state with any assurance the relationship between rivalry rate and figure-ground contrast, it will be necessary first to investigate the changes in performance which occur with practice and also the particular effect of massed practice.

SUMMARY AND CONCLUSIONS

The results of a previous experiment suggested that the rate of alternation in binocular rivalry may be related to the degree of figure-ground contrast of stereoscopic fields. In order to investigate this relationship five stereograms, differing in degree of figure-ground contrast, were presented to five *S*s. No reliable systematic relationship was found although a trend in this direction is indicated. It was suggested that the higher intersubject variability found in this experiment over the previous one was due to the influence of massing of the trials.

(Received for priority publication August 25, 1952)

REFERENCES

1. ALEXANDER, L. T. The influence of figure-ground relationships in binocular rivalry. *J. exp. Psychol.*, 1951, **41**, 376–381.
2. FISHER, R. A., & YATES, F. *Statistical tables for biological, agricultural, and medical research* (3rd Ed.). Edinburgh: Oliver and Boyd, 1948.

DISCRIMINATION OF TACTUAL STIMULI

HERBERT J. BAUER

Loyola University, Chicago

The rapid, errorless identification and manipulation of control levers, knobs, and push buttons has, for some years now, been an aim of equipment designers. The problem which this aim presents is of particular interest to those involved in the design of military and naval apparatus, specifically as related to radio, sonar, radar, fire control, aircraft, and similar devices (**2, 3, 4, 6**). In an effort to accomplish the efficient operation of such machines recent investigators have based their work on the principle that ". . . every little facility can help make the job faster, more accurate, and free from infrequent but important mistakes" (**3**, p. 308).

To date five general methods of facilitating the operation of controls have been coding by (*a*) location, (*b*) color, (*c*) size, (*d*) mode of actuation, and (*e*) shape. Each of these has, in some measure, contributed to the more rapid learning and more efficient operation of controls. The writer's aim in his current research is to investigate the possibility of coding controls in terms of variations in *textures* of the control's surface. That such application of the tactual sense seems reasonable is indicated by the results of investigations reported by Major (**5**), Zigler and Barrett (**7**), and Zigler and Northrup (**8**). More recently Austin and Sleight (**1**) have indicated some of the potentialities of the tactual sense. The present paper reports only the results of the preliminary research that was done in an effort to determine which, of a number of randomly designed tactual stimuli, are sufficiently discriminable to be employed for texture coding typical control lever surfaces.

Method

Materials.—The ten textures used in this experiment were randomly designed. Each stimulus was a $2 \times 2 \times \frac{1}{2}$-in. block of wood with one side textured. These were as follows: (A) polished wood surface with parallel grooves $\frac{1}{4} \times \frac{1}{4}$ in. spaced on $\frac{1}{4}$-in. centers; (B) $\frac{1}{4}$-in. hemispheres raised approximately $\frac{1}{8}$ in. from the surface, and lying adjacent to one another in a diagonal pattern across the wooden block; (C) surfaced with Norton Mohawk abrasive flint paper, grit No. 2; (D) surfaced with Jewel abrasive cloth, grit No. 3/0; (E) polished surface of the wood block; (F) polished wood surface with parallel grooves $1/32 \times 1/32$ in. on $\frac{1}{4}$-in. centers; (G) polished wood surface with grooves $1/32 \times 1/32$ in. running perpendicularly to each other across the square spaced on $\frac{1}{4}$-in. centers; (H) wood surface "roughed up" by pitting with very small gouges randomly made; (I) polished wood surface with 3/16-in. diameter holes on $\frac{1}{4}$-in. centers, about $\frac{1}{4}$ in. deep; (J) polished wood surface with irregularly spaced wavy grooves $\frac{1}{4} \times \frac{1}{4}$ in.

Procedure.—In *Method I* the blindfolded *S* rested the palm of his hand on a support but kept his index, middle, and fourth fingers out straight, over the edge of a table. The first stimulus was raised to come in contact with *S*'s extended finger tips; no sliding movement of stimulus or fingers was permitted. This was the experimental stimulus which was to be remembered by *S*. Immediately after this, *S* was presented, in the same way, with all ten stimuli, including the one he had examined as "experimental." The exposure time in each presentation was 1 sec. The *S* was instructed to remember the experimental stimulus and to indicate, as soon as it was presented in the series, that he was again in contact with it. In other words, *S* had to recall and identify the experimental stimulus during the serial presentation of all ten stimuli, thereby discriminating between it and the nine other textures.

Each of the ten stimuli served, in turn, as the experimental stimulus which had to be identified when presented in series with the others.

The procedure in *Method II* was identical with that in Method I except that *S* was permitted to draw his three fingers simultaneously

across each surface once. Exposure time was 1 sec.

The procedure in *Method III* was the same as that of Method II, except that *S* wore a thin, cotton glove tightly fitted over the exploring fingers.

Because preliminary work indicated that variations in the amount of pressure in contact influenced correct identification, a relatively equal pressure in all exposures to the stimuli was maintained. This was accomplished by mounting the textures on a spring-controlled plunger. *S* held his fingers rigid so that the pressure could be determined by *E* bringing the contact to 10-oz. pressure as shown by a guide line painted on the shaft of the plunger.

The relative position between *S* and the stimuli was held constant throughout the series so that the directionality of the stimulus patterns was the same for all *S*s.

The order of presenting the experimental stimuli to each *S* and the position of the experimental stimuli in the selection series were determined by the method of systematized randomization. In this way variables in selection position, the length of time any stimulus had to be remembered, and the number and nature of stimuli preceding the experimental stimulus in the selection series were equated.

Subjects.—All *S*s were male college students. No *S* was used for more than one complete series, during which each of the ten surfaces was used as experimental stimulus only once. Interrogation after the experiment indicated that the *S*s were not aware of this fact. A total of 40 *S*s was used: 25 for Method I, 10 for Method II, and 5 for Method III.

Results

Since the criterion for this experiment was the ability of *S*s to perceive, and later correctly identify, a given tactual stimulus, the results of the three methods were recorded in terms of the number of times a stimulus was correctly identified during the serial presentations. It was also considered of interest to note the sources of errors where incorrect identifications were made, i.e., which stimuli were confused with which other stimuli. Finally, it was desired to determine the relative merits for identification of each of the stimuli.

Table 1 shows the extent of confusion and failure for the three methods.

TABLE 1

Frequency with Which Experimental Stimuli Were Matched with Various Stimuli

Matched Stimulus	Experimental Stimulus A	B	C	D	E	F	G	H	I	J
					Method I ($N = 25 \times 10$)					
A	12									
B	3	17								
C	3	5	16							
D	—	—	3	5						
E	—	—	1	7	9					
F	—	—	—	1	16	2				
G	—	2	5	1	2	5	10			
H	—	—	1	20	9	10	5	2		
I	—	—	3	5	2	3	2	7	12	
J	14	1	5	2	—	—	6	—	2	8
*	—	—	1	—	—	1	1	—	2	1
	A	B	C	D	E	F	G	H	I	J
					Method II ($N = 10 \times 10$)					
A	8									
B	—	10								
C	—	—	10							
D	—	—	—	9						
E	—	—	—	—	6					
F	—	—	—	—	6	5				
G	—	—	—	—	—	—	7			
H	—	—	—	—	1	1	2	7		
I	—	—	—	—	—	—	—	1	8	
J	3	—	—	—	—	—	1	—	3	7
*	—	—	—	1	—	1	1	1	1	—
	A	B	C	D	E	F	G	H	I	J
					Method III ($N = 5 \times 10$)					
A	5									
B	—	4								
C	—	—	5							
D	—	—	—	5						
E	—	—	—	—	2					
F	—	—	—	—	6	2				
G	—	—	—	—	—	—	4			
H	—	—	—	—	—	1	—	4		
I	—	—	—	—	—	—	1	—	3	
J	—	1	—	—	—	—	—	—	3	4
*	—	—	—	—	—	—	—	—	—	—

* Indicates failure of *S* to select any stimulus as "just like" the experimental stimulus.

The rows and columns of the table represent the letter names for each of the stimuli. The intersections of the rows and columns indicate the number of times the stimuli were confused with each other. For example, with Method I, Stimulus A was identified correctly 12 times, reported as Stimulus B 3 times, as Stimulus C 3 times, and as Stimulus J 14 times. With Method II, Stimulus B was correctly identified in all presentations. It was

TABLE 2

PERCENTAGE OF CORRECT IDENTIFICATIONS

Stimulus	Methods		
	I	II	III
A	48	80	100
B	68	100	80
C	64	100	100
D	20	90	100
E	36	60	40
F	8	50	40
G	40	70	80
H	8	70	80
I	48	80	60
J	32	70	80

never confused with any other stimulus when presented under these conditions.

Table 2 shows the percentage of correct identifications of the stimuli by each method. Table 3 shows the ranking, based on percentage of correct responses, for all stimuli and all methods.

Table 1 shows that, regardless of the method of presentation and inspection, certain of the stimuli were never confused with certain others. Typical is Stimulus A which was never confused with D, E, F, G, H, or I. Stimulus J was never confused with E, F, or H. Stimulus C was never confused with F.

TABLE 3

RANK OF STIMULI BASED ON PERCENTAGE OF CORRECT IDENTIFICATIONS

Stimulus	Method			
	I	II	III	All
A	3.5	4.5	2.0	3.3
B	1.0	1.5	5.5	2.7
C	2.0	1.5	2.0	1.8
D	8.0	3.0	2.0	4.3
E	6.0	9.0	9.5	8.2
F	9.5	10.0	9.5	9.7
G	5.0	7.0	5.5	5.8
H	9.5	7.0	5.5	7.3
I	3.5	4.5	8.0	5.3
J	7.0	7.0	5.5	6.5

Conversely we may note that certain of the stimuli were frequently confused with each other. Outstanding is the confusion between A and J as well as the confusion between E and F. Stimulus H was readily confused with D, E, F, and G.

In the over-all, Stimulus C was the most stable (Tables 2 and 3), appearing in second rank for two of the three methods, and with the highest mean rank (1.8). Stimulus B, first in rank for two of the three methods, is second in the over-all ranking (mean rank 2.7), while the third best texture, with a mean rank of 3.3, was Stimulus A.

It must be realized that the merits of the stimuli reported here are relative. The percentages and ranks reported are based on the given series of stimuli as a whole. The reported figures could readily be modified by changing the combinations of the number of stimuli used. If, for example, we were to eliminate Stimulus J, the percentage of correct identifications for Stimulus A would be greatly increased. The disparity of the percentages of correct identifications and the relative rankings is a function of the particular structure and content of the given series, and not merely of the nature of the individual stimulus pattern taken independently.

The introduction of movement did, in all cases, improve *S*'s ability to identify the stimuli (Table 2).

Statistical evaluation of the results obtained in this experiment indicates that the accuracy of identification of all stimuli by Method II was better than would be expected by chance alone. All of these percentages are significant at the .01 level or better. By Method I, all percentages except those for stimuli D, F, and H were significant at least at the .01 level. The small number of *S*s used for

Method III did not warrant statistical analysis of the percentages obtained.

Discussion

Undoubtedly much of the confusion in the identification of certain stimuli arises from the fact that they are not sufficiently unique to be perceived as distinctly different when the exposure time to the stimulus is as limited as it was in this experiment. This is especially true for Method I wherein the basic determiner of the percept is related to the two-point threshold. The confusion between pairs A–J, D–H, and E–F seems to point out this fact clearly, for the stimulus patterns presented to *S* by touch alone are quite similar for the pairs mentioned.

It is apparent that permitting *S* to draw his finger tips over the stimuli presented him with additional cues useful in perceiving and discriminating between stimuli. The intermittent stimulation of the receptors occasioned by moving them across the stimuli probably enhanced the perception of more definite form of the patterns involved. The reduction of the amount of confusion between A and J with the introduction of movement may well have been caused in this manner. The continuous stimulation of certain areas of the fingertips when moving across A is readily distinguishable from the intermittent stimulation of the same points when moving across J.

Kinesthetic sensations, too, seem to have played a contributing role in the discrimination of stimuli. The finger tips moving across the smooth surface of E would encounter little resistance, but when moving across C (with or without the glove) a comparatively greater amount of resistance is presented. The possibility of such a cue facilitating discrimination is suggested by the reduction in the amount of confusion between D and H in Methods II and III, as compared with Method I.

With the introduction of movement, *Ss* (according to their reports) seem to have been aided by two other cues: vibration and sound. Certainly the movement of the fingers across rough surfaces such as C, B, or J would set them into some sort of vibratory motion. The variations in these vibration patterns apparently facilitated discrimination.

Moving the fingers over some of the textures gave rise to audible sounds; others did not. Unfortunately, the presence of this phenomenon as a cue for identification did not come to the attention of *E* until it was pointed out by the last two *Ss* (Method III).

Introspections were called for after each *S* completed his series of identifications. *Ss* were interrogated regarding general difficulties and impressions. Only two *Ss* reported the use of sound cues in the interviews. Admittedly, this lack of control is an error in the experimental design and should have been controlled by the use of a masking sound.

The data clearly indicate that permitting *S* even minor degrees of movement over the stimuli greatly enhances his ability to perceive and discriminate between the various surfaces presented. The *precise* nature of the cues thus obtained are, however, not so evident.

One might readily presume that the introduction of a barrier between the finger tips and the stimulus (as in Method III) would introduce a reduction of the accuracy of identification. Tables 1 and 2 show that this was not the case. The scatter of errors for Method III is less than for Methods I or II. It is possible that the greater resistance brought about by the catching of the cotton glove on the projec-

tions of some of the surfaces actually aided discrimination by the introduction of kinesthetic cues. It was observed, for instance, that in the cases of the smoother surfaced stimuli most *Ss* would draw their fingers across the entire length of the block before making judgments. In the cases of the coarser stimuli (such as C and D) only a small portion of the square would be tactually examined before a decision was made as to the identify or nonidentity of the stimuli in question. Stimuli D and H stand out as highly improved with the introduction of movement and the glove.

Application.—That it is possible for *Ss* to discriminate, without visual cues, between various surface textures has been demonstrated. Whether or not this ability may be used in practical applications will depend, at least in part, upon the demands of the situation in which it is to be used. The question of the practicality of texture coding of industrial or military equipment must be determined through further research. Such an investigation is currently in progress.

Summary

This experiment was designed to determine which, of a number of randomly designed tactual stimuli, are sufficiently discriminable to be feasibly employed for the purpose of texture coding typical control knobs, levers, and push buttons. Ten different surface textures were used. Three different methods of serial exploration were permitted.

It was found that simple, 1-sec. tactual contact with the textures was not as efficient as 1-sec. exploration permitting movement. Even the introduction of a thin cotton glove between the stimuli and the finger tips did, in some instances, facilitate identification.

The stimuli were ranked according to per cent correct identifications. The cues used by *Ss* were found to be based primarily in touch, perception of form, kinesthesis, vibration, and sound.

Accuracy of discrimination between textures was, in the majority of cases, significantly better than would be expected by chance alone.

It was deemed feasible to investigate the extent to which textured surfaces, such as those used in the experiment, might be utilized in coding control surfaces of levers, knobs, and push buttons on military and naval devices.

(Received for priority publication August 28, 1952)

References

1. Austin, T. R., & Sleight, R. B. Accuracy of tactual discrimination of letters, numerals, and geometric forms. *J. exp. Psychol.*, 1952, **43**, 239–247.
2. Chapanis, A., *et al. Lectures on men and machines.* Baltimore: Systems Research Laboratory, 1947.
3. Chapanis, A., *et al. Applied experimental psychology.* New York: Wiley, 1949.
4. Jenkins, W. O. The tactual discrimination of shapes for coding aircraft-type controls. In P. M. Fitts (Ed.), *Psychology of equipment design.* Report No. 19, Army Air Forces Aviation Psychology Program Research Reports. Washington: U. S. Govt. Print. Off., 1947. Pp. 199–205.
5. Major, D. R. Cutaneous perception of form. *Amer. J. Psychol.*, 1898, **10**, 143–147.
6. Weitz, J. The coding of aircraft control knobs. In P. M. Fitts (Ed.), *Psychology of equipment design.* Report No. 19, Army Air Forces Aviation Psychology Program Research Reports. Washington: U. S. Govt. Print. Off., 1947. Pp. 187–198.
7. Zigler, M. J., & Barrett, R. A further contribution to the tactual perception of form. *J. exp. Psychol.*, 1927, **10**, 184–192.
8. Zigler, M. J., & Northrup, K. M. The tactual perception of form. *Amer. J. Psychol.*, 1926, **37**, 391–397.

VISUAL FACTORS IN THE PERCEPTION OF VERTICALITY

CECIL W. MANN

Tulane University

The judgment of verticality is a function of the joint action and interaction of visual and proprioceptive stimulus variables. Under conditions of visibility the intact organism makes use of the vertical and horizontal visual framework of the world about him in his judgments of the upright, and in the absence of visual stimulation can use his proprioceptive equipment to assume and maintain vertical posture. In some of his activities, man has to make perceptual judgments when visual and proprioceptive variables are in conflict. This is true of the military pilot during and following unusual maneuvers in the air.

The data of laboratory experiments on the judgment of visual and postural verticality have been variously interpreted in favor of the dominance either of the visual or the proprioceptive variables. Gibson (3) has recently suggested a resolution of this problem in terms of the interaction of visual and proprioceptive stimulation. When the main lines of the retinal image are parallel to the direction of the force of gravity, the visual-proprioceptive stimulus variables are consistent. When these directions are not physically parallel, the situation is one of conflicting or discrepant variables. This condition may be produced by rotating the visual image either by optical or mechanical means or by adding a centrifugal component to that of gravity. He points out that discrepant conditions are not produced merely by lateral tilting of the body; in this situation the direction of gravity has not been altered, albeit it passes through the body at an unusual angle.

> The theory which emerges from these considerations is something like this. In the case of reciprocal visual-proprioceptive stimulation, the coupled variables combine to form an *invariant* resultant which is in correspondence with the objective direction of gravity and which provides the stimulus for a univocal impression of the vertical. This is why the ordinary visual scene continues to look upright when one inclines his head and thereby rotates his retinal image. In the case of discrepant visual-proprioceptive stimulation, or conflicting cues, the coupled variables do not yield a resultant. Their values do not correspond in the usual fashion, and the resulting perception is ambiguous or equivocal. The subject is instigated to perceive two different verticals at the same time. The organism is forced to choose between them (3, p. 79).

Passey (5) and Passey and Ray (6) found that subjects tilted laterally at inclinations up to 20° in right and left quadrants, and required to set a visual target to the visual vertical made mean constant errors which ranged from 3.6° at 5° body tilt to 4.6° at 20° body tilt. In their experiments, the equipment consisted of the Navy tilting room. The chair in which *S* was seated and the room (which could be tilted independently of the chair) were inclined laterally at preset conditions, and *S*, remaining in the tilted position, was required to bring the room to the true vertical. It will be noted that these conditions do not satisfy a test of Gibson's theory.

[1] This report is one of a series of investigations conducted jointly with the U. S. Naval School of Aviation Medicine under Contract N7-onr-434, T. 0.1, with the Office of Naval Research in cooperation with the Bureau of Medicine and Surgery and the Tulane University of Louisiana. Thanks are due Mr. Paul E. Thetford for his assistance in collecting the data of this experiment.

When *S* opens his eyes, he may be faced with a discrepant condition, i.e., the visual framework (room) is tilted and not parallel to the direction of gravity. When, however, he commences to bring the room towards the gravitational vertical, the discrepancy is reduced, and when the room is at the gravitational vertical, the condition is one of consistency. In order to test Gibson's hypothesis, it is necessary to have *S* place the visual target in alignment with the gravitational vertical while he remains in the preset consistent or discrepant condition. This can be done by introducing a visual target against a tiltable visual framework as background.

In a recent experiment, Boring (**2**) set up these conditions in the Tulane tilting chair. In the absence of visual stimulation, *S* was required to set a luminescent target-rod to the gravitational vertical against visual frameworks consisting of (*a*) a luminescent straight line 14 in. in length and (*b*) a window pattern, 14 × 12 in. The tilt of *S*'s body and the tilt of the visual framework were maintained while he made his judgments of the vertical. Boring obtained mean average errors and mean constant errors of the same order as those found by Passey, and by Passey and Ray. There were no differences between the errors made under consistent and under discrepant conditions. It was suggested in this report that the visual frameworks used did not produce visual stimulation strong enough to conflict with the proprioceptive stimulation, and that therefore the visual stimulus was ignored by *S* in making his judgments.

In this connection, it should be pointed out that there may be conflict between visual variables as well as between visual and proprioceptive variables. For instance, an upright telephone pole seen against a steep hillside may appear tilted in the opposite direction. The main visual lines are the direction of the hillside slope. Under this strong visual framework, it is the tilt of the pole rather than that of the hillside that is affected. This recalls Wertheimer's experiment (**7**) in which he had *S*s look through a tube at a mirror image of a room inclined so that the verticals and horizontals were tilted at 45°. He reports that at first the room appeared tilted, but eventually the tilt disappeared and the room appeared upright. Reporting a similar experiment, Asch and Witkin (**1**) state that "an appreciable number of subjects accepted the tilted scene as completely upright or as almost completely upright." Gibson and Mowrer (**4**) repeated the experiment and report that following the observation of the tilted room for a minute or two the room appeared more "natural" but still remained tilted throughout the observation. In commenting on their results, they emphasize the point that the illusion of uprightness was maintained only when *S* was able to project himself into the mirror image, and conclude that Wertheimer was reporting a cognitive process—the decrease in the strangeness and unreality of the mirror image.

In order to supplement Boring's experiment, it appears necessary to increase the strength of the visual stimulus at least to the point at which *S* is able to identify himself with the visual framework. Only then will there be real conflict between the visual and the proprioceptive variables. When conflict does occur, one would predict greater errors for discrepant than for consistent conditions, and errors for consistent conditions no greater than those made under normal conditions of observations.

Procedure

In this experiment the visual framework consisted of the Navy tilting room. Consistent and discrepant conditions were imposed by suitable preset conditions of the tilting room and tilting chair, and *S* was required to set a visual target-rod to the direction of the gravitational vertical while maintaining the preset consistent or discrepant condition. The Navy tilting room, described elsewhere (5), was modified to provide a visual target-rod at a distance of 5 ft. at *S*'s eye level, and 1 ft. from the rear wall of the tilting room. The target consisted of a selsyn-controlled rod, 1 in. wide and 30 in. in length, adjustable by *S* in the chair or by *E* in the control room. The room and the chair were adjustable by *E* to the nearest degree, the target-rod was adjustable to an accuracy of .04°. This was possible by the inclusion of a suitable gear-train between the target-rod and the recording device. The *S* was requested to close his eyes during the adjustment of the inclinations of the equipment, which was at all times well illuminated. The room adjustments were 30°L, 10°L, 0°, 10°R, and 30°R; the chair adjustments were 30°L, 0°, and 10°R. This made a total of 15 conditions presented to *S* in random order. The target-rod was set an equal number of times to 10°L and 10°R of the true gravitational vertical, arranged in random order. Each of the nine *S*s made 18 judgments under each condition; the judgments were spread over six sessions each of 45 trials for each *S*.

Nine male *S*s between ages 18 and 20 yr. were used. They were without previous experience in this type of experiment or equipment. On entering the test room, *S* was given the opportunity of watching the independent tilting movements of the room, the chair, and the target-rod. He was instructed that his problem was to adjust the target-rod to the direction of the gravitational vertical while the chair or the room were in preset positions. The direction of gravitational force was described as being parallel to the walls of the laboratory or at right angles to the floor of the hangar in which the tilt chair was located. At the beginning of each session, emphasis was placed upon the nature of the problem, and the meaning of vertical in terms of the walls and floor of the hangar.

This experiment was designed to test Gibson's hypothesis of consistent and discrepant variables. There are other problems which could be investigated in this arrangement. Under certain conditions of room and chair tilt, *S* will have a conflict of verticals from which to choose. He may feel his body tilted and attempt to make an adjustment of the target-rod to compensate for this amount of tilt; he may regard the tilted room as vertical and believe that *E* wishes him to align the target-rod to the direction of the walls of the titled room; or he may try to adjust the target-rod to the direction of the walls of the laboratory. Were the problem one of investigating what and how people choose under a conflict of conditions, it is possible that naive *S*s would serve the purpose. Since, however, the study was concerned with the judgment of the gravitational vertical, it seemed desirable that each *S* be instructed and practiced in the selected criterion of the investigation.

TABLE 1

Mean Constant Errors in the Judgment of the Vertical under Consistent and Discrepant Visual–Proprioceptive Stimulus Conditions

Chair Position	Discrepant		Consistent	Discrepant	
	Room 30°L	Room 10°L	Room 0°	Room 10°R	Room 30°R
30°L	2.58L	4.22L	0.45R	6.90R	16.35R
0°	9.61L	1.59L	0.08R	4.04R	11.71R
10°R	13.69L	1.29L	0.88L	2.91R	11.93R

Results

The mean constant errors and mean average errors in the setting of the target-rod to the gravitational vertical are shown in Table 1 and Table 2, respectively. The means in each cell in each table are based upon 18 judgments for each of nine *S*s, a total of 162 judgments in each condition. The mean average errors made with the

TABLE 2

Mean Average Errors in the Judgment of the Vertical under Consistent and Discrepant Visual–Proprioceptive Stimulus Conditions

Chair Position	Discrepant		Consistent	Discrepant	
	Room 30°L	Room 10°L	Room 0°	Room 10°R	Room 30°R
30°L	17.39	7.50	3.08	8.86	18.21
0°	12.27	4.95	0.72	5.37	13.74
10°R	15.32	6.90	1.75	5.63	14.50

room tilted at 10°L and 10°R are significantly different at the .10 level from those made with room at 0°; those made at 30° inclinations are significantly different at the .01 level. The mean constant errors show a trend towards significance at 10° room tilts; and the errors at 30° tilts are significant. It is to be noted that the mean constant errors are predominantly in the direction of room tilt.

The mean constant error of 2.58°L for the cell in which the chair is 30°L and the room is 30°L is difficult to explain. This value is strongly influenced by the results of one *S* whose constant error in this cell only was strongly to the right instead of to the left as would be expected. All his other judgments were predominantly in the direction of room tilt. There is no justification for the omission of his results, but if this were done, the mean constant error for the group would be 12.84°L, a mean constant error more in keeping with the mean average error, and with the rest of the mean constant errors in the table.

Discussion

The results of this experiment are far more decisive than those obtained in the experiment of Boring. Under consistent conditions—with visual frameworks upright and subject tilted at either 30°L or 10°R—the mean errors are not significantly greater than those made by the subject seated in an upright position. These results are in accord with Gibson's suggestion that the tilt of the visual field produced by body tilt will be compensated by the corresponding stimulation of the statocysts to produce an invariant resultant.

The prediction that greatest errors would occur under discrepant conditions is supported by the results. Even when the body is upright, the 30°L and 30°R inclinations of the visual framework produce significant errors in judgment in the direction of the visual tilt. The largest errors occur when the visual framework is inclined at an angle of 30°. It should also be noted that the magnitude of the error of judgment, in every case, is far less than the angle of inclination of the visual framework. Even under such strongly conflicting conditions, the judgment is the resultant of interaction of visual and proprioceptive variables.

These results seem to add weight to Boring's suggestion that discrepant conditions will not be effective in influencing the judgment until the subject can identify himself with the visual framework. This identification does not take place apparently when the framework is a short straight line or a simple window pattern of the size used by Boring. In his experiment the conflicting elements were the relatively weak visual framework and the relatively strong proprioceptive stimulation, and judgments were made predominantly in terms of the latter.

The present experimental results combined with those of Boring would appear to make necessary a modification of Gibson's hypothesis. Basically, there is agreement with the proposal that under consistent conditions, there is compensatory visual-proprioceptive interaction, and with the prediction that consistent conditions will produce small errors in the judgment of the vertical. The statement of the quality of the judgments under discrepant conditions needs modification in terms of the effective value of the visual stimulation. When the visual stimulation is increased to the point at which the subject can identify himself with the visual framework, there is conflict between the visual stimulation produced by the framework and the proprioceptive

stimulation produced by gravitational force acting on the statocysts. The subject then has two verticals to deal with: one produced by his identification with the visual framework, the other by the gravitational forces acting on his body. Under this condition, and depending upon the relative strengths of the conflicting forces, the judgments may be compromised.

Summary

1. The experiment was designed to test the hypothesis stated by Gibson to the effect that the error in the perception of the vertical under consistent visual-proprioceptive stimulation will be less than that made under discrepant conditions. Consistent visual-proprioceptive conditions occur when the vertical lines of the retinal field are parallel to the direction of the gravitational vertical; discrepant conditions are produced when these directions are not parallel.

2. Nine *S*s were tested in 15 experimental conditions of room and body tilt. In each condition *S* was required to set a target-rod in alignment with the gravitational vertical.

3. An analysis of the average and constant errors gives support to Gibson's hypothesis.

4. Comparing the results of this experiment with those of a previous experiment by Boring, one concludes that when visual stimulation is increased to the point at which *S* can identify himself with the visual framework, there is conflict between the visual stimulation produced by the framework and the proprioceptive stimulation produced by the gravitational force acting on the statocysts.

5. The hypothesis offered by Gibson needs a qualifying statement to the effect that in the judgment of the vertical reliance will be placed upon the proprioceptive stimulus variables until the visual factors in the visual-proprioceptive variables become sufficiently effective to offer conflict with the proprioceptive variables. When this happens, a compromise judgment will be made, the judgment lying somewhere between the inclination of the visual field and the gravitational vertical.

(Received for priority publication August 28, 1952)

References

1. Asch, S. E., & Witkin, H. A. Studies in space perception. I. Perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, **38**, 325–337.
2. Boring, R. O. The effect of visual stimulus variables upon the perception of the visual vertical. U. S. Naval Sch. Aviat. Med. and Tulane Univer. Joint Report No. 28, 1952.
3. Gibson, J. J. The relation between visual and postural determinants of the phenomenal vertical. U. S. Navy Dep.: *ONR Sympos. on Spatial Orientation*. 1950, 77–80.
4. Gibson, J. J., & Mowrer, O. H. Determinants of the perceived vertical and horizontal. *Psychol. Rev.*, 1938, **45**, 300–323.
5. Passey, G. E. The perception of the vertical: IX. Adjustment of the visual vertical from various magnitudes of body tilt. U. S. Naval Sch. Aviat. Med. and Tulane Univer. Joint Report No. 15, 1950.
6. Passey, G. E., & Ray, J. T. The perception of the vertical: X. Adaptation effects in the adjustment of the visual vertical. U. S. Naval Sch. Aviat. Med. and Tulane Univer. Joint Report, No. 17, 1950.
7. Wertheimer, M. Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.*, 1912, **61**, 161–265.

ERRATUM

CLYDE E. NOBLE

Perceptual and Motor Skills Research Laboratory, Human Resources Research Center

In my recent paper, "The Role of Stimulus Meaning (m) in Serial Verbal Learning" (*J. exp. Psychol.*, 1952, **43**, 437–446) a decimal-placing error has been detected in the within-cells σ^2 estimate of Table 2. The F ratios for column ($\bar{m}$) and row (rest-nonrest) differences should read 3.67 and 5.23, respectively, with $p < .05$ in each case. With this correction, retention just following the 7/12 criterion is shown to increase significantly with increasing $\bar{m}$-value, while the color-naming "rest" period results in significant forgetting: .86 of an item. Interpretations on pages 442 and 445 should be altered appropriately. Thus freed of error, the analysis indicates an even stronger case for the absence of reminiscence under these conditions.

(Received September 22, 1952)

Classics Among
PSYCHOLOGICAL MONOGRAPHS

Thorndike, E. L. The Mental Life of the Monkey. 1899, #15. *$.50*

Carr, Harvey. Visual Illusion of Movement During Eye Closure. 1905, #31. *$1.25*

Watson, John B. Kinaesthetic and Organic Sensations: Their Role in the Reactions of the White Rat to the Maze. 1907, #33, *$1.00*

Shepherd, William T. Some Mental Processes of the Rhesus Monkey. 1910, #52. *$.75*

Franz, Shepard Ivory and Lafora, Gonzalo R. On the Functions of the Cerebrum: The Occipital Lobes. 1911, #56. *$1.25*

Fernberger, Samuel W. On the Relation of the Methods of Just Perceptible Differences and Constant Stimuli. 1912, #61. *$1.00*

Boring, Edwin G. Learning in Dementia Praecox. 1913, #63. *$1.00*

Langfeld, Herbert S. On the Psychophysiology of a Prolonged Fast. 1914, #71. *$.75*

Franz, Shepard Ivory. I. Symptomological Differences Associated with Similar Cerebral Lesions in the Insane. II. Variations in Distribution of the Motor Centers. 1915, #81. *$1.50*

Peckstein, Louis Augustus. Whole versus Part Methods in Motor Learning. 1917, #99. *$.75*

Kjerstad, Conrad L. The Form of the Learning Curves for Memory. 1919, #116. *$1.35*

Tolman, Edward C. Retroactive Inhibition as Affected by Conditions of Learning. 1918, #107. *$.75*

MANY OF THE EARLY MONOGRAPHS ARE OUT-OF-PRINT.
ONLY A LIMITED QUANTITY OF THE ABOVE
NUMBERS ARE AVAILABLE.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1333 Sixteenth Street N.W., Washington 6, D. C.

10 1472